Roberto Torretti es profesor emérito de la Universidad de Puerto Rico y miembro titular del Institut International de Philosophie y de la Académie Internationale de Philosophie des Sciences. Ha sido *Humboldt Dozentenstipendiat*, *Guggenheim Fellow* (dos veces) y *Fellow* del Center for Philosophy of Science de Pittsburgh. Reconocido mundialmente por sus análisis de la relatividad, el espaciotiempo y su geometría, ha publicado numerosos artículos y también los siguientes libros: *Diccionario de lógica y filosofía de la ciencia* (Madrid, 2002; en colaboración con Jesús Mosterín), *The Philosophy of Physics* (Cambridge, 1999), *El Paraíso de Cantor* (Santiago, 1998), *Relativity and Geometry* (Nueva York, 1996), *La geometría del universo* (Mérida, 1994), *Variedad en la razón* (San Juan, 1992; en colaboración con Carla Cordua) y *Creative Understanding* (Chicago, 1990).

Roberto Torretti

Relatividad y espaciotiempo

RIL® editores,

por la

bibliodiversidad

Roberto Torretti

Relatividad y espaciotiempo



530.1 Torretti Edwards, Roberto
T Relatividad y espaciotiempo / Roberto Torretti Edwards. -- 1ª ed. -- Santiago: RIL® editores, 2003.

272 p. -- 21 cm.

1 RELATIVIDAD (FÍSICA). 2 ESPACIO Y TIEMPO. I Torretti Edwards, Roberto. II t.

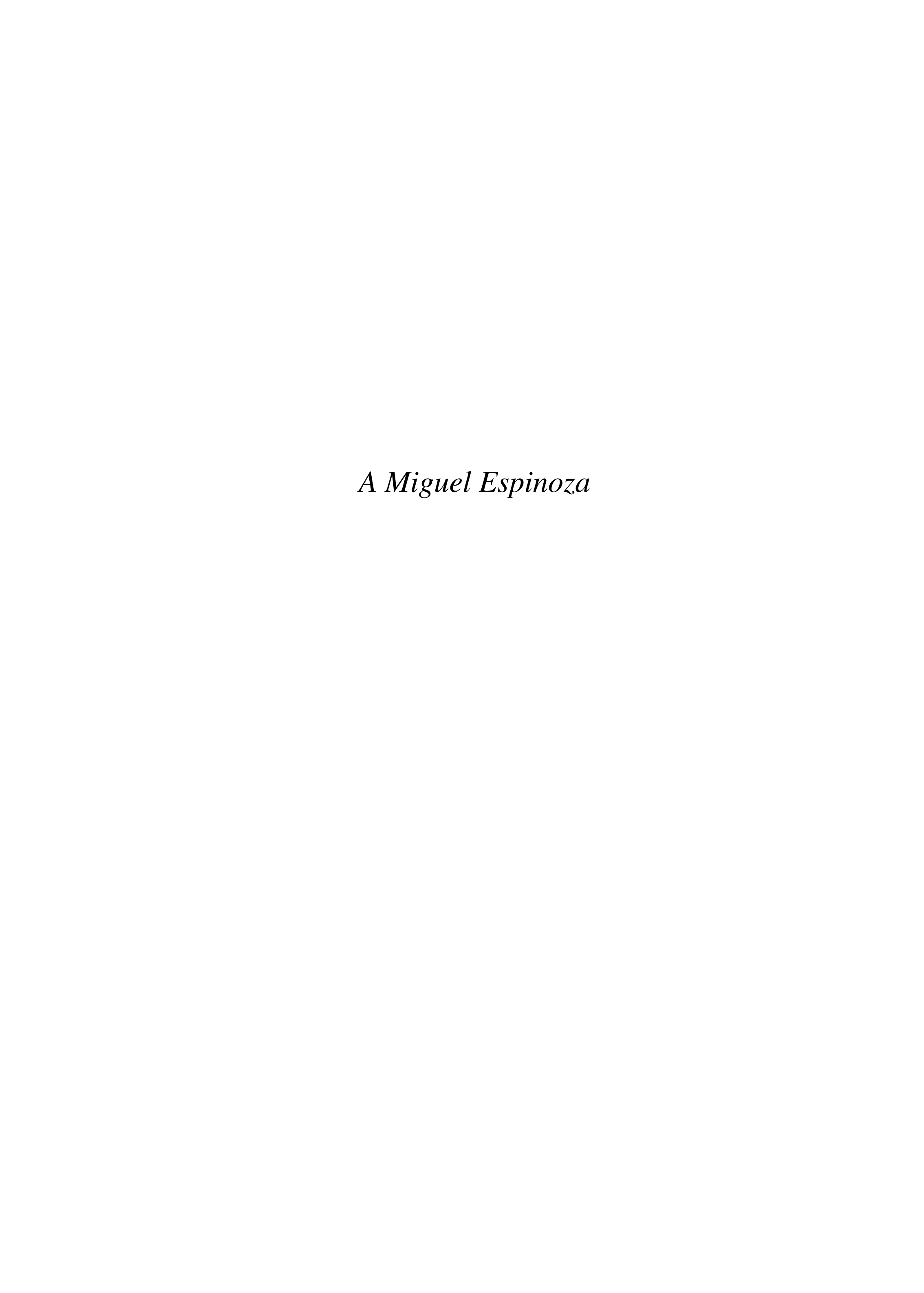
ISBN 956-284-283-5

RIL® editores
El Vergel 2882, of. 11, Providencia
Santiago de Chile
Tel. (56-2) 2238100 - Fax 2254269
ril@rileditores.com / www.rileditores.com

Baldomero Fernández Moreno 1217
Buenos Aires, Argentina
Tel. (54-11) 4432-2840

Composición e impresión: RIL® editores
Diseño de portada: Cristián Silva L.

Impreso en Chile - *Printed in Chile*



A Miguel Espinoza

Sumario

Prólogo

En el título de este libro y conforme al uso ordinario, la palabra 'relatividad' se refiere a dos teorías de Albert Einstein que no hay que confundir, una de 1905 y otra de 1915, a las que él denominó *la teoría especial* y *la teoría general de la relatividad*, respectivamente, en su libro popular de 1917. Para mayor brevedad, llamaré *relatividad especial* a la primera y *relatividad general* a la segunda. La diferencia y las relaciones entre ambas teorías se explican en el capítulo 1, así como el origen de sus nombres; pero no está demás señalar desde ya que, si no estuviesen tan arraigados, sería justo reasignarlos. Pues la teoría *especial* preside la descripción y explicación de *todas* las formas conocidas de interacción física, al menos *localmente*, en el ámbito de los laboratorios donde se las estudia; mientras que la teoría llamada *general* es *solo* una teoría de la gravitación, la mejor que tenemos para concebir y describir *globalmente* esta interacción que es la más débil de todas, pero también la de más largo alcance. Como se irá viendo a lo largo del libro, sería más apropiado llamarlas respectivamente la *teoría local* y la *teoría global* del *espaciotiempo*; pero Hermann Minkowski, que abrió la perspectiva que autoriza estas denominaciones y que probablemente no habría tenido inconveniente en adoptarlas, murió a los 44 años en 1909, cuando no había aún dos teorías que distinguir. También el concepto de espaciotiempo y su papel central en ambas teorías de Einstein se explican en el capítulo 1, pero se irá haciendo más concreto en los capítulos siguientes.

Mientras redactaba el capítulo 1 —prometido a Anna Estany para el tomo de la *Enciclopedia Iberoamericana de Filosofía* sobre filosofía de las ciencias exactas, naturales y sociales editado por ella— y veía crecer el manuscrito más allá de la extensión prescrita, me vino la idea de reunirlo en un libro con versiones castellanas de cuatro trabajos sobre el pensamiento de Einstein y sus continuadores que escribí en inglés en 1999 y 2000. Tres de ellos aparecieron ya en las publicaciones que se indican en la p. 263, mientras que el cuarto —dividido aquí entre los capítulos 5 y 7— fue distribuido como lectura

suplementaria a los partícipes de la escuela de verano sobre filosofía y cosmología patrocinada por la Universidad Meléndez y Pelayo de Santander en julio de 2000. Los cuatro trabajos fueron escritos a invitación de distintas personas que propusieron los temas respectivos, aunque, presumo, adivinando que me interesaría tratarlos. Pueden por eso leerse independientemente y en cualquier orden, aunque, si se sigue el orden en que aquí aparecen y que fue el mismo en que fueron escritos, se percibirá que es el más natural.

El capítulo 2, que aquí titulo "El orden del tiempo", se refiere a las aportaciones de las dos teorías de Einstein a este tema, la relatividad especial con su célebre crítica de la simultaneidad y sus consecuencias, la relatividad general en cuanto admite formas de orden y desorden completamente ajenas al tiempo lineal de nuestra tradición judeocristiana. Lo escribí para el libro sobre el tiempo que editó Jeremy Butterfield con ocasión del centenario de la British Academy y el tema me fue indicado por él.

El capítulo 3, "Gravedad y curvatura", se basa en la charla que di a comienzos de mayo de 1999 en el III Simposio de Seven Pines (Wisconsin, Estados Unidos). El tema lo fijó el organizador, Roger Stuewer, que quería una presentación introductoria y didáctica que precediera a las ponencias más especializadas de John Earman, Robert Wald y Robert Geroch.

También el capítulo 4, "Modelos del mundo", fue planeado como charla, para el simposio sobre espaciotiempo organizado por Michael Redhead que tuvo lugar en la London School of Economics a fines de mayo de 1999. No pude asistir, pero el texto que aquí traduzco apareció en un número especial de *Studies in the History and Philosophy of Modern Physics*, junto con las ponencias efectivamente presentadas en dicha ocasión.

Jesús Mosterín, que dirigió —junto a Miguel Boyer— la escuela de verano sobre filosofía y cosmología de Santander, deslindó muy exactamente el tema del capítulo 5, "Los universos de Friedmann: derivando la geometría". Se trataba de exponer "la estructura matemática del modelo estándar del *Big Bang* de un modo razonablemente preciso, claro y explícito, dando todas las definiciones requeridas y señalando los pasos principales de las deducciones". Debía dirigirme a un lector "interesado en la filosofía de la cosmología (… filó-

sofo, físico, matemático, periodista,…), no un especialista, pero tampoco un ignorante. Algunos lectores se saltarán las fórmulas, por incapaces de entenderlas, y otros por ya conocerlas, pero en cualquier caso muchos aprenderán el tema leyéndolo…" Cito sus palabras porque describen al lector a quien me gustaría llegar con el libro entero.

El capítulo 5 es con mucho el más técnico de todos. Para agilizar la exposición me pareció oportuno relegar las definiciones a un apéndice, que aquí va en el capítulo 7, ya que su contenido les sirve también a los seis primeros. Hay un índice alfabético separado (pp. 265 y ss.) de los términos definidos en este apéndice, y cuando se los utiliza por primera vez en uno de los capítulos precedentes van seguidos de una pequeña estrella, así☆; si el término en cuestión es una frase, va encerrada entre dos de estos signos ☆como aquí☆.

Dado que las teorías de Einstein introdujeron en la física del siglo XX las ideas sobre geometría desarrolladas en el siglo anterior por Felix Klein y Bernhard Riemann, me pareció oportuno suplementar los cinco capítulos sobre relatividad y espaciotiempo con una traducción castellana del artículo sobre la geometría del siglo XIX que John Norton me instó a escribir para la *Stanford Encyclopedia of Philosophy* y que está disponible desde 1999 en la página web http://plato.stanford.edu/entries/geometry-19th/. Va en el capítulo 6, el menos técnico de todos.

Los "suplementos" que forman el capítulo 8 son tres ensayos breves o notas largas sobre asuntos a que se alude en tres de los capítulos.

Como cabía esperar de trabajos sobre temas muy próximos entre sí, escritos casi al mismo tiempo para publicaciones independientes, contenían muchas repeticiones. He eliminado las más sencillas —definiciones, citas textuales, datos históricos—, reemplazándolas con referencias cruzadas. Pero no he tocado las más complejas, estimando que no deja de ser instructiva la reiteración de las mismas ideas en otro contexto y con otro enfoque.

He uniformado las referencias bibliográficas. En el texto y las notas cito cada obra, como es habitual, por el apellido del autor y la fecha de publicación (o por una sigla, cuando la fecha de la edición utilizada es poco significativa). Cuando cito dos o más obras de un mismo autor publicadas en el mismo año, las distingo con letras añadi-

das a la fecha; en general, dichas letras son arbitrarias, pero en el caso de las obras de Einstein anteriores a 1922 uso las mismas con que se las señala en sus *Collected Papers*. La lista de obras citadas que figura en las pp. 247 y ss. da las fichas completas.

He relegado las notas al final del libro, en vez de colocarlas al pie de página. Los encabezados de la sección de notas indican las páginas en que aparecen las llamadas correspondientes. Recomiendo leer primero el texto de corrido y sin consultar las notas, a menos que la llamada acompañe a un término técnico cuyo significado el lector desconoce y probablemente hallará explicado en la nota.

Al índice de términos matemáticos definidos en el capítulo 7, sigue un índice analítico general que abarca los otros seis capítulos. Recomiendo acudir a él para ubicar la explicación de términos no asociados a una nota ni marcados con una estrellita. También es aconsejable consultar el diccionario de Mosterín y Torretti (2002).

Las ecuaciones están numeradas consecutivamente dentro de cada capítulo. Aquellas que figuran en las notas llevan el número que les correspondería si la nota estuviese insertada en el texto en el lugar de la llamada correspondiente. Cuando se hace referencia a una ecuación de otro capítulo, el número de este precede al de la ecuación. Por ejemplo, "la ecuación (7.39)" designa la ecuación (39) del capítulo 7.

No está demás advertir que mi traducción de los capítulos que escribí en inglés no siempre se atiene a la letra de los originales; en cambio, es probablemente más fiel al espíritu e intención del autor que la mayoría de las traducciones.

Agradezco a Anna Estany, Jeremy Butterfield, Roger Stuewer, Michael Redhead, Jesús Mosterín y John Norton el impulso inicial para escribir estos trabajos. Pero el ánimo decidido y persistente sin el cual no podría nunca haberlos realizado se lo debo a Carla Cordua. Doy gracias, pues, con mucha alegría, a Carla y a la Moira que hila nuestras cosmolíneas.

Santiago de Chile, 11 de octubre de 2002.

1

Relatividad y espaciotiempo

1.1 Espacio y tiempo en la física newtoniana

La física moderna, matemática y experimental, fundada por Newton y perfeccionada por Euler, Lagrange, Laplace, Fourier, Hamilton, Helmholtz, Kelvin y Maxwell, entre otros, concibe la naturaleza como una cantidad fija de materia que se mueve eternamente en el espacio infinito. Newton le señaló su tarea así: "Partiendo de los fenómenos del movimiento investigar las fuerzas de la naturaleza, y luego, partiendo de estas fuerzas, demostrar los demás fenómenos".[1] Se trata, por cierto, de investigar las fuerzas existentes, que causan cambios efectivos en el movimiento real —y no solo aparente— de los cuerpos. Este consiste, según Newton, en su desplazamiento en el espacio "absoluto, verdadero y matemático" en el curso del tiempo "absoluto, verdadero y matemático". Por otra parte, como Newton mismo anota, nuestra descripción precisa de los fenómenos del movimiento tiene que referirlos al espacio "relativo, aparente y vulgar" determinado por un cuerpo rígido que se adopta como estándar (por ejemplo, el triedro formado por las paredes y el piso de un laboratorio) y al tiempo "relativo, aparente y vulgar" medido por un movimiento reputado regular (por ejemplo, la rotación de la tierra).

El abismo entre la verdad absoluta y matemática que se trata de conocer y la apariencia relativa y vulgar que nos es accesible sería infranqueable —y el proyecto de la física moderna habría nacido muerto— de no mediar las dos circunstancias siguientes: 1º La primera "ley del movimiento" postulada por Newton —el llamado *principio de inercia*[2]— garantiza la uniformidad de ciertos movimientos que, por lo tanto, constituyen relojes naturales. 2º Las leyes newtonianas del movimiento implican que las relaciones entre movimientos en el espacio absoluto son formalmente idénticas a las relaciones entre

movimientos referidos a un cuerpo rígido que se mueve conforme a la primera ley,[3] de modo que basta estudiar estas para averiguar las leyes y las fuerzas de que dependen aquellas. Hacia fines del siglo xix, Carl Neumann, Ludwig Lange y James Thomson aprovecharon estas dos características de la teoría de Newton para eliminar el espacio y el tiempo absolutos de entre los conceptos fundamentales de la física. En vez del tiempo absoluto, Neumann (1870) propone la *escala inercial de tiempo*, en la cual dos lapsos de tiempo t_1 y t_2 son iguales si y solo si una partícula libre recorre en ambos distancias iguales. En vez del espacio absoluto, Thomson (1884) y Lange (1885) —independientemente— proponen el *marco inercial de referencia*, en cuyo espacio tres partículas libres que salen de un punto en tres direcciones diferentes trazan líneas rectas.

Muchos filósofos se resisten a aceptar que la fantasía metafísica pueda tan fácilmente ceder su puesto en la base misma de la física a lo que a fin de cuentas no son más que estipulaciones convencionales. Pero no hay que perder de vista que se trata aquí de convenciones *bene fundatae*, acordadas en vista del comportamiento efectivo de las cosas: una vez que se ha elegido como estándar del tiempo el movimiento de *una* partícula libre, *cualquier otra* partícula libre se mueve a velocidad constante según la escala inercial adoptada; una vez que se ha fijado un marco inercial de referencia mediante *tres* partículas libres, *cualquier otra* partícula libre se mueve con respecto a él en línea recta; más aún, *cualquier otro* conjunto análogo de tres partículas libres determina otro marco inercial equivalente al primero.

Con todo, las nuevas definiciones dejan una importante laguna sin llenar. Un observador que viaje con una partícula libre puede utilizar hitos equidistantes en su camino para determinar, conforme a la escala inercial de tiempo, la hora en que pasa frente a cada uno de ellos; pero esta es, en cada ocasión, la hora ahí, en el lugar donde se encuentra. Newton escribió una vez que "cada instante del tiempo se difunde de suyo por todo el espacio universal";[4] pero quien no comparta esta oscura creencia tendrá que reconocer, con James Thomson, que la hora establecida en un lugar solo puede difundirse a otros por medio de señales. Sin embargo, parecería que cualquier método de señalización producirá aproximadamente los mismos resultados, y

—teniendo en cuenta la inevitable imprecisión de las observaciones— Thomson se dio por satisfecho con esta apariencia.[5]

1.2 Relatividad del movimiento inercial y velocidad de la luz

Es claro entonces que, diga Newton lo que diga, su mecánica no dispone de un estándar absoluto de movimiento y reposo: conforme a sus principios, no es posible distinguir entre el reposo real y el movimiento uniforme y rectilíneo. Pero el siglo XIX creyó hallar un sucedáneo del espacio absoluto en el éter luminífero, el medio a la vez rígido y elástico que se suponía portador de las ondas de luz. El persistente fracaso de las hipótesis propuestas para explicar su naturaleza no impidió a los científicos perseverar en la postulación de su existencia con la tozuda fe del carbonero. Con la electrodinámica de Maxwell, que incorpora la luz en un vasto espectro de ondas electromagnéticas de todas las frecuencias concebibles, parece confirmarse definitivamente la necesidad de referir todos los movimientos al éter. En las ecuaciones de Maxwell figura una constante c, con la dimensión de una velocidad, que resulta ser igual a la velocidad de la luz en el vacío. ¿A qué sino al éter puede estar referida esta velocidad? Por otra parte, la velocidad de una partícula eléctricamente cargada contribuye a determinar la magnitud y dirección de la fuerza que ejerce sobre ella el campo magnético. Ello supone, obviamente, que dicha velocidad esté referida a un estándar de reposo, el cual, al parecer, no podría ser otro que el éter, en que dicha fuerza se asienta. Un primer paso necesario para acceder al estándar y hacer de él nuestro marco universal de referencia sería medir la velocidad con que nuestro planeta se desplaza a través del éter en cada momento.

Llamemos V a la velocidad de la tierra en el éter. Si el sol reposa en el éter, V oscila alrededor de los 30.000 m/s; si también el sol se mueve, el valor medio de V podría ser mucho mayor. Sin embargo, como Einstein recuerda al comienzo de su artículo sobre la electrodinámica de los cuerpos en movimiento (1905*r*), todos los intentos de

medir la velocidad V daban un resultado negativo. Sea c la velocidad relativa al éter de una señal luminosa (de aquí en adelante, debe entenderse que cualquier señal luminosa aludida en este libro, salvo expresa indicación contraria, se propaga en el vacío). Hasta 1881, los experimentos para medir la velocidad de la tierra en el éter solo eran capaces de revelar efectos del orden del cociente V/c. El resultado negativo de los experimentos se explicaba entonces fácilmente, pues, según la teoría ondulatoria de la luz, los términos de este orden de magnitud se cancelan entre sí. El interferómetro inventado ese año por A. A. Michelson era sensible a efectos del orden de $(V/c)^2$, pero no los puso de manifiesto (Michelson y Morley 1887). Otro experimento (Michelson y Morley 1886) forzaba a descartar la hipótesis —adelantada por Stokes— de que el éter es arrastrado por la atmósfera terrestre. G. F. Fitzgerald y H. A. Lorentz conjeturaron entonces que, al moverse a través del éter con velocidad V, los cuerpos rígidos se contraen de tal modo que su dimensión paralela al movimiento se multiplica por $\sqrt{1-(V/c)^2} < 1$. La diferencia resultante entre las longitudes de los dos brazos perpendiculares de que consta el interferómetro de Michelson explicaría entonces el resultado negativo de su experimento. Una lectura atenta de los escritos más maduros de Lorentz indica que contienen implícita una tesis análoga sobre la marcha de los relojes: cuando se mueven a través del éter el tiempo medido con ellos entre dos eventos dados se dilata en la proporción $1{:}\sqrt{1-(V/c)^2} > 1$.

Einstein (1905*r*) aborda esta dificultad de un modo completamente distinto. En lugar de recurrir a hipótesis físicas sobre los relojes y las varas de medir cuya justificación requeriría un conocimiento de la microestructura de la materia que nadie tenía en ese entonces, postula dos principios que, según él, son un epítome de los fenómenos. Conforme al *principio de relatividad*, si $\mathcal{R}$ es un sistema de referencia con respecto al cual cualquier partícula libre reposa o se mueve uniformemente en línea recta, esto es, si $\mathcal{R}$ es —como diré en adelante— un *marco inercial de referencia,* y $\mathcal{R}'$ es un sistema de referencia que se mueve uniformemente en línea recta respecto a $\mathcal{R}$, $\mathcal{R}'$ es también un marco inercial de referencia y *ningún experimento de física permite*

establecer que uno de ellos reposa mientras que el otro se mueve. Conforme al *principio de la constancia de la velocidad de la luz*, una señal luminosa se propaga siempre con la misma velocidad c respecto a un marco inercial dado —y, por ende, según el principio de relatividad, respecto a todos ellos—, *no importa cual sea la velocidad de la fuente emisora.*

A primera vista, estos dos principios parecen totalmente incompatibles. Consideremos dos marcos inerciales $\mathcal{R}$ y $\mathcal{R}'$, tales que $\mathcal{R}'$ se mueve respecto a $\mathcal{R}$ en una dirección fija con velocidad uniforme igual a v. Supongamos que una fuente en reposo en el punto O de $\mathcal{R}$ emite en todas direcciones un pulso de luz en el instante en que O coincide con el punto O' de $\mathcal{R}'$. Trascurrido un lapso de tiempo T después de ese instante, el frente de la onda luminosa emitida llenará una esfera con centro en O y radio $r = cT$. En ese momento, el punto O' está separado de O por una distancia igual a vT. Sin embargo, según el principio de relatividad, el frente de la onda también debe llenar en ese preciso momento una esfera con radio $r = cT$ pero con su centro en O'.[6] Evidentemente, esto no puede ser.

Einstein mostró que la incompatibilidad entre ambos principios es solo aparente. El razonamiento presentado en el párrafo anterior presupone que, si la luz se propaga en $\mathcal{R}$ y $\mathcal{R}'$ con la misma velocidad c y la onda luminosa salió de O cuando este punto de $\mathcal{R}$ coincidía con el punto O' de $\mathcal{R}'$, el frente de la onda llegará simultáneamente a todos los puntos de $\mathcal{R}$ que distan ct de O y a todos los puntos de $\mathcal{R}'$ que distan ct de O'. Pero este supuesto ni siquiera tiene sentido para nosotros a menos que tengamos un criterio para juzgar si sucesos que ocurren en distintos puntos del espacio son o no simultáneos entre sí. Como bien vio James Thomson, la física newtoniana nunca fijó tal criterio. Einstein (1905*r*, §1) propone dos. El primero es muy sencillo: la recepción E, aquí y ahora, de una señal luminosa se reputa simultánea con la emisión F de esa señal desde otro lugar. Einstein rechaza este criterio porque depende del lugar de recepción adoptado como punto de referencia.[7] Adopta en cambio el segundo criterio, también bastante simple. Con arreglo a éste, un evento[8] E ocurrido en un punto O de un marco inercial $\mathcal{R}$ se reputa simultáneo con un evento F ocurrido en otro punto P de $\mathcal{R}$ si y solo si F coincide con la recepción y reemisión en

P de una señal luminosa *S* emitida en *O* y el tiempo trascurrido entre la emisión de *S* y el evento *E* es igual al tiempo trascurrido entre *E* y la recepción en *O* de la señal *S* después que rebotó en *P*. Por lo tanto, si la coordenada temporal de la emisión de *S* en *O* es t_1 y la de la recepción de *S* en *O* es t_2, la coordenada temporal de *F* es ½$(t_2 - t_1)$. Este criterio de simultaneidad entre eventos distantes, aunque convencional, está respaldado por la naturaleza de las cosas. Como Einstein se adelanta a señalar, el criterio es viable porque no depende ni del lugar desde donde se emite la señal *S* ni del momento en que se la emite. Si aplicamos el método descrito para sincronizar mediante la señal *S* un conjunto de buenos relojes que reposan en diferentes puntos del marco inercial $\mathcal{R}$, esos relojes seguirán sincronizados —dentro de límites de error admisibles— si el mismo método vuelve a aplicarse usando otra señal emitida más tarde desde el mismo lugar o desde otro lugar diferente de $\mathcal{R}$. Llamo *tiempo de Einstein adaptado a* $\mathcal{R}$ a la coordenada temporal definida de este modo. Dos tiempos de Einstein adaptados a un mismo marco inercial pueden diferir a lo sumo por una traslación (que desplaza el instante al que se asigna la coordenada cero) y un cambio en la unidad de medida; por lo tanto, ambos determinan la misma partición de todos los eventos en clases de eventos simultáneos. Por tanto, hablando de la simultaneidad entre eventos relativa a un marco inercial dado, es lícito referirse a *el* tiempo de Einstein adaptado a ese marco.

Atendamos nuevamente al caso del pulso de luz emitido desde *O* en $\mathcal{R}$ cuando este punto coincide con el punto *O*′ del marco $\mathcal{R}'$ que se mueve respecto a $\mathcal{R}$ con velocidad constante. Una breve reflexión bastará para convencernos de que la llegada del frente de la onda a cada punto de la esfera de radio *r* y centro *O* fija en $\mathcal{R}$, que es simultánea según el tiempo de Einstein adaptado a $\mathcal{R}$, no puede ser simultánea según el tiempo de Einstein adaptado a $\mathcal{R}'$, mientras que su llegada a cada punto de la esfera de radio *r* y centro *O*′, simultánea según el tiempo de Einstein adaptado a $\mathcal{R}'$, no puede serlo según el tiempo de Einstein adaptado a $\mathcal{R}$. Puede entonces darse el caso de que tres eventos *A*, *B* y *C* sean simultáneos según el tiempo de Einstein adaptado a $\mathcal{R}$, mientras que *A* precede a *B* y *B* precede a *C* según el tiempo de Einstein adaptado a $\mathcal{R}'$. Esta comprobación disuelve la

apariencia de que el principio de relatividad y el principio de la constancia de la velocidad de la luz son incompatibles. Pero indica también que su afirmación conjunta, aunque lógicamente consistente y acorde con la experiencia, impone por otra parte un cambio grandísimo en la concepción ordinaria del tiempo: la *simultaneidad* y la *sucesión temporal* de los eventos pasan a ser *relativas al marco inercial* al que esté referida su descripción.

Este resultado tiene numerosas consecuencias a primera vista paradojales. Consideraré solo unas pocas. Sean, como antes, $\mathcal{R}$ y $\mathcal{R}'$ dos marcos inerciales y v la magnitud de la velocidad relativa con que cada cual se mueve relativamente al otro. Sean t y t' tiempos de Einstein adaptados respectivamente a $\mathcal{R}$ y $\mathcal{R}'$, que miden el tiempo en la misma unidad y asignan ambos el 0 a cierto evento E que ocurre en el punto O de $\mathcal{R}$ cuando este punto coincide con el punto O' de $\mathcal{R}'$. Sea E^* un evento que ocurre más tarde en O cuando este coincide con otro punto P' de $\mathcal{R}'$. Entonces la coordenada $t(E^*)$ que t le asigna a E^* típicamente difiere de la coordenada $t'(E^*)$ que le asigna t'. Se puede demostrar que $t'(E^*) = t(E^*) \Big/ \sqrt{1-\left(v/c\right)^2}$. Supongamos que E y E^* son el comienzo y el fin de un proceso que trascurre en O. Es claro que este proceso dura menos en el tiempo de Einstein adaptado al marco de referencia $\mathcal{R}$ en que O reposa que en el adaptado al marco de referencia $\mathcal{R}'$ en que O se mueve. Más aún, cualquiera que sea la duración del proceso según el tiempo de $\mathcal{R}$, su duración según el tiempo de $\mathcal{R}'$ crece sin límite a medida que la velocidad con que O se mueve en este marco de referencia se aproxima a la velocidad de la luz. Un resultado análogo afecta a las distancias. Sean O y P dos puntos que reposan en $\mathcal{R}$. En aras de la simplicidad, supondremos que la recta que los une es paralela al movimiento de $\mathcal{R}$ en $\mathcal{R}'$. La distancia entre O y P en $\mathcal{R}$ se determina, por ejemplo, colocando varas de medir en reposo a lo largo de esa recta. Pero ¿cómo medir la distancia entre O y P en $\mathcal{R}'$? El método siguiente parece razonable: marcar dos posiciones simultáneas de O y P en $\mathcal{R}'$ y medir la distancia entre ellas con varas de medir que reposan en $\mathcal{R}'$. Se puede demostrar entonces que, si λ es la distancia entre O y P en $\mathcal{R}$ y λ' es la distancia, así determinada, entre O y P en $\mathcal{R}'$,

$$\lambda' = \lambda\sqrt{1-\left(v/c\right)^2} \tag{1}$$

A medida que *v* se aproxima a la velocidad de la luz, la distancia λ′ tiende a 0.

El lector atento habrá advertido que, si *O* y *P* son los extremos de uno de los brazos del interferómetro de Michelson, la longitud de ese brazo, medida en un marco de referencia en que este brazo se mueve, en la dirección de *O* a *P*, con velocidad *V*, es menor que la longitud del mismo brazo en reposo en exactamente la proporción prevista por Fitzgerald y Lorentz. Pero desde el nuevo punto de vista inaugurado por Einstein, no tenemos que habérnoslas con un proceso material de contracción, causado por una fuerza desconocida, sino solo con una diferencia de valor entre dos cantidades diferentes, a saber, la distancia permanente entre los extremos del brazo en el marco inercial en que reposa y la distancia entre dos posiciones simultáneas de esos extremos en el marco de referencia en que el brazo se mueve longitudinalmente con velocidad *V*. Como estas dos posiciones no son simultáneas en el otro marco de referencia, esas dos cantidades no tienen por qué ser iguales. También habrá advertido el lector que la proporción $1{:}\sqrt{1-\left(V/c\right)^2}$ en que —según está implícito en Lorentz— se retardaría la marcha de un reloj que se mueve en el éter con velocidad *V* es exactamente igual a la proporción $t'(E^*){:}t(E^*)$ entre las duraciones de un mismo proceso según el tiempo t' adaptado a un marco inercial en que el lugar donde ocurre el proceso se moviera con velocidad *V*, y según el tiempo *t* adaptado a un marco inercial en que dicho lugar reposa. Pero bajo la perspectiva abierta por Einstein, no se trata de una alteración en el funcionamiento de los relojes causada por su movimiento, sino de una diferencia en las cantidades que los relojes están llamados a medir y que suponemos que en cada caso miden bien.[9]

La afirmación conjunta de los dos principios de Einstein permite acoger sin cambios las leyes de la electrodinámica clásica referidas indiferentemente a cualquier marco inercial (y no solo, como se pensaba, al marco de referencia en que reposa el éter). Pero obliga a

introducir modificaciones en las leyes clásicas de la mecánica. La nueva mecánica relativista hace prácticamente las mismas predicciones que la newtoniana cuando se trata de cuerpos que se mueven lentamente respecto al marco inercial de referencia elegido, pero sus predicciones son muy diferentes cuando las velocidades en juego se acercan a c. Además la reforma de la mecánica alteró profundamente algunos conceptos. Mientras que en la mecánica newtoniana la inercia —esto es, la resistencia que un cuerpo material opone a sus cambios de velocidad— tiene un valor fijo igual a lo que Newton llamaba la *cantidad de materia* contenida en ese cuerpo y que más tarde se llamó su *masa*, en la mecánica relativista dicha inercia depende de la misma velocidad v que se trata de cambiar, de acuerdo con la fórmula:

$$m(v) = \frac{m_0}{\sqrt{1-\left(v/c\right)^2}} \qquad (2)$$

donde m_0 mide la resistencia que el cuerpo opondría a cualquier intento de sacarlo del reposo, esto es, su inercia relativa al marco inercial en que reposa momentáneamente. De aquí se sigue inmediatamente que la inercia de los cuerpos supera todo límite asignable cuando v se acerca a c, un fenómeno verificado hasta la saciedad en los aceleradores de partículas y que suele citarse como la prueba más palpable del acierto de Einstein. En tales condiciones es obviamente imposible que una partícula masiva alcance la velocidad de la luz.

1.3 El espaciotiempo de Minkowski

Años más tarde, en varios escritos didácticos, Einstein señaló una asimetría en el tratamiento del espacio y del tiempo por la física newtoniana. Todos sus forjadores habrían reconocido sin vacilar que el *lugar* en que ocurren dos eventos sucesivos solo puede identificarse —o distinguirse— relativamente a un marco de referencia;[10] pero ninguno de ellos pensó que el *instante* en que ocurren dos eventos distantes tampoco se puede identificar —o distinguir— salvo con

respecto a tal marco de referencia. Al tomar conciencia de esta similitud entre la coincidencia temporal o simultaneidad entre eventos y su coincidencia espacial o isotopía, Einstein abrió paso a la concepción del espacio y el tiempo combinados en un *espaciotiempo*, que comúnmente asociamos a su teoría de la relatividad. Sin embargo, no fue él quien la introdujo, sino Hermann Minkowski, que había sido uno de sus profesores de matemáticas en el Politécnico de Zürich y a la sazón era colega de Felix Klein y David Hilbert en la Universidad de Göttingen.[11]

A quien contempla en la noche desde una torre de vigilancia junto a un cruce de autopistas los vehículos que transitan velozmente por ambas y pasan de una a otra por un trébol le parece mucho más natural concebir lo que ve como un solo haz de historias continuas que como una infinitud de instantáneas tomadas sucesivamente a un recipiente atemporal de contenido cambiante. Pero en 1900, aunque el automóvil y el cinematógrafo ya se habían inventado, la concepción física del acontecer seguía dominada por la experiencia común de cuerpos casi rígidos y aparentemente inmóviles —muros, techos, suelo— entre los que reposamos o nos movemos lentamente. De ahí en parte la insistencia en utilizar, para la representación del acontecer físico, dos continuos distintos de una y de tres dimensiones, respectivamente, y la resistencia a combinarlos en uno solo de cuatro, como propuso más de alguien en los siglos XVII y XVIII. Tal propuesta no se hizo esperar después que matemáticos y físicos adoptaron la representación de los lugares del espacio mediante coordenadas introducida por Descartes (1637). Un sistema de coordenadas cartesianas x, y, z asigna en forma exclusiva a cada punto espacial P un triple de números reales $\langle x(P),y(P),z(P)\rangle$ iguales a las distancias desde P a los tres planos ortogonales de un triedro de referencia, multiplicadas por 1 o por –1 según a qué lado del plano correspondiente se encuentre P. Como la práctica de asignar números a sucesos estaba establecida desde la antigüedad no solo entre los historiadores, sino también y con mayor precisión entre los astrónomos, tiene que haber parecido obvio que un evento E que ocurre en un lugar P puede representarse inequívocamente mediante cuatro números reales, a saber, la coordenada temporal $t(E)$, y las coordenadas espaciales $x(E) = x(P)$, $y(E) = y(P)$, y $z(E)$

= $z(P)$. Pero en el contexto de la física newtoniana este cuádruplo numérico se deja analizar naturalmente y de un solo modo en dos componentes, el tiempo $t(E)$ y las coordenadas de posición $\langle x(E), y(E), z(E)\rangle$, como ahora veremos.

En el resto de la §1.3 consideraremos sistemas alternativos de coordenadas, t, x, y, z y t', x', y', z', que representan el tiempo medido en segundos y las distancias medidas en metros.[12] Requerimos además que las coordenadas t y t' crezcan ambas en la dirección del futuro y que las x, y, z y x', y', z' sean coordenadas cartesianas definidas con respecto a marcos inerciales de referencia $\mathcal{R}$ y $\mathcal{R}'$, respectivamente, y que crecen en la dirección del pulgar (las x), el índice (las y) y el cordial (las z) de una mano derecha con estos tres dedos abiertos formando ángulos rectos entre ellos. De un sistema de coordenadas cartesianas que satisfaga este último requisito diré que está *orientado por la derecha*. Bajo estos supuestos, la *transformación de coordenadas* que determina las coordenadas t', x', y', z' de cualquier evento E en función de las coordenadas t, x, y, z del mismo evento, pertenece a una de las tres clases descritas a continuación o resulta de la aplicación sucesiva de dos o más transformaciones de estas clases.

I. $\mathcal{R}$ y $\mathcal{R}'$ reposan el uno respecto al otro y pueden tratarse como idénticos. El segundo sistema de coordenadas se obtiene por *traslación* del origen del primero. Entonces:

$$
\begin{aligned}
t' &= t + a_t \\
x' &= x + a_x \\
y' &= y + a_y \\
z' &= z + a_z
\end{aligned}
\qquad (3)
$$

donde las constantes a_t, a_y, a_y y a_z son los parámetros característicos de la traslación.

II. $\mathcal{R}$ y $\mathcal{R}'$ reposan el uno respecto al otro y los orígenes temporal y espacial de ambos sistemas coinciden.[13] El segundo se obtiene por *rotación* del primero en torno a su origen. Entonces:

$$
\begin{aligned}
t' &= t \\
x' &= a_{11}x + a_{12}y + a_{13}z \\
y' &= a_{21}x + a_{22}y + a_{23}z \\
z' &= a_{31}x + a_{32}y + a_{33}z
\end{aligned}
\tag{4}
$$

donde los parámetros de la rotación a_{ij} cumplen la condición $a_{ij} = a_{ji}$ $(1 \leq i, j \leq 3)$.

III. El origen temporal de ambos sistemas coincide. En el instante $t = t' = 0$ coinciden también sus orígenes espaciales y los ejes de las x, las y y las z apuntan respectivamente en las mismas direcciones que los ejes de las x', las y' y las z'. $\mathcal{R}'$ se mueve respecto a $\mathcal{R}$ recorriendo v_x m/s en la dirección positiva de las x. Hasta principios del siglo xx se daba por descontado que en este caso

$$
\begin{aligned}
t' &= t \\
x' &= x - v_x t \\
y' &= y \\
z' &= z
\end{aligned}
\tag{5}
$$

Diremos que esta es una *transformación cinemática de Galileo* con parámetro v_x.[14] Obsérvese que, en virtud de las ecuaciones (5), una señal luminosa emitida en $\mathcal{R}$ desde el origen en la dirección positiva del eje de las x llega en un segundo al punto $x = c$; como en el mismo tiempo llega al punto $x' = c - v_x$ en $\mathcal{R}'$, es claro que, bajo los supuestos en que se basa la transformación cinemática de Galileo, la velocidad de la luz no puede ser la misma en todos los sistemas inerciales.

La aplicación sucesiva de dos o más transformaciones de uno de los tres tipos descritos es una transformación de ese tipo. Por lo tanto, cada uno de estos tipos de transformaciones constituye un *grupo*.[15] Combinando traslaciones del grupo (I) con rotaciones del grupo (II) se genera el grupo de los *movimientos euclidianos.* Esta denominación se explica así: si P y Q son dos puntos de este espacio y x, y, z es un sistema de coordenadas cartesianas como los descritos arriba,

la distancia entre P y Q es igual a la raíz cuadrada de $(x(P)-x(Q))^2$ + $(y(P)-y(Q))^2$ + $(z(P)-z(Q))^2$ y para cualquier otro sistema x', y', z' relacionado con x, y, z por una transformación del tipo (I) o del tipo (II) o por cualquier combinación de tales transformaciones se cumple que

$$\begin{aligned}(x(P)-x(Q))^2+((y(P)-y(Q))^2+(z(P)-z(Q))^2 =\\ (x'(P)-x'(Q))^2+(y'(P)-y'(Q))^2+(z'(P)-z'(Q))^2\end{aligned} \tag{6}$$

Después que Klein (1872) propuso identificar a cada geometría con el estudio de los *invariantes* de un determinado grupo de transformaciones, esto es, de las propiedades y relaciones que estas preservan, la geometría euclídea quedó asociada justamente con el grupo generado por las traslaciones y rotaciones (véase la §6.3).

Al comienzo de su exposición clásica sobre el espaciotiempo, Minkowski (1909) observa que, mientras el grupo de los movimientos euclidianos es reconocido como un carácter fundamental del espacio, el grupo de las transformaciones puramente cinemáticas casi no recibe atención. Por eso, no se ha pensado en el grupo compuesto de ambos, a pesar de que este grupo total es justamente el que "nos da que pensar" (p. 104). Este cambio de perspectiva obedece por cierto a una consecuencia que ya Einstein (1905*r*) había derivado de sus dos principios:[16] si afirmamos conjuntamente la relatividad del movimiento inercial y la constancia de la velocidad de la luz, y si t y t' son tiempos de Einstein adaptados respectivamente a $\mathcal{R}$ y $\mathcal{R}'$, entonces una transformación cinemática del tipo (III) no satisface las ecuaciones (5), sino las ecuaciones siguientes:

$$\begin{aligned}t' &= \frac{t-\left(v_x x/c^2\right)}{\sqrt{1-\left(v_x/c\right)^2}}\\ x' &= \frac{x-v_x t}{\sqrt{1-\left(v_x/c\right)^2}}\\ y' &= y\\ z' &= z\end{aligned} \tag{7}$$

donde c denota el número de metros que una señal luminosa recorre en un segundo (es decir, 299.792.458).[17] Diremos que esta es una *transformación cinemática de Lorentz* dependiente del parámetro v_x. Apliquémosla a una señal luminosa emitida en $\mathcal{R}$ desde el origen en la dirección positiva del eje de las x. Si en $\mathcal{R}$ esta señal, como suponemos, recorre c = 299.792.458 metros por segundo, cuando $t = 1$ estará llegando a $x = c$; pero esto equivale a decir que llega a $x' = (c - v_x)(1 - v_x{}^2 c^{-2})^{-1/2}$ justo cuando $t' = (1 - (v_x c / c^2))(1 - v_x{}^2 c^{-2})^{-1/2} = c^{-1}(c - v_x)(1 - v_x{}^2 c^{-2})^{-1/2}$; por lo tanto, en $\mathcal{R}'$ la señal recorre x'/t' = c = 299.792.458 metros por segundo, igual que en $\mathcal{R}$. El grupo generado por las rotaciones (II) y las transformaciones que satisfacen las ecuaciones (7) es el *grupo* (homogéneo, conectado) *de Lorentz*. El generado por este grupo y las traslaciones (I) es *el grupo* (conectado) *de Poincaré* o grupo inhomogéneo de Lorentz.[18]

Una rápida inspección de las ecuaciones (5) y (7) constata la diferencia señalada por Einstein. El paso de $\mathcal{R}$ a $\mathcal{R}'$ conforme a la transformación (5) de Galileo rompe las relaciones de isotopía entre eventos pero deja intactas las de simultaneidad. En cambio, la transformación (7) de Lorentz también afecta a la simultaneidad; además, el lapso de tiempo $|t'(E_1) - t'(E_2)|$ entre dos eventos E_1 y E_2 tales que $t(E_1) = t(E_2)$ aumenta con la velocidad v_x y la diferencia $|(x(E_1)) - (x(E_2))|$. Por otro lado, la transformación (7) —al igual que la (5) pero siguiendo una regla diferente— induce un cambio en la distancia entre los lugares donde ocurren E_1 y E_2: en general, la distancia $(x(E_1) - x(E_2))^2 + (y(E_1) - y(E_2))^2 + (z(E_1) - z(E_2))^2$ difiere de $(x'(E_1) - x'(E_2))^2 + (y'(E_1) - y'(E_2))^2 + (z'(E_1) - z'(E_2))^2$. Esta variación en la distancia que separa los dos eventos se combina precisamemente con la variación del intervalo temporal entre ellos para asegurar que:

$$
\begin{aligned}
&\left(t(E_1)-t(E_2)\right)^2 \\
&\quad -c^{-2}\left(\left(x(E_1)-x(E_2)\right)^2+\left(y(E_1)-y(E_2)\right)^2+\left(z(E_1)-z(E_2)\right)^2\right) \\
&\quad = \left(t'(E_1)-t'(E_2)\right)^2 \\
&\quad -c^{-2}\left(\left(x'(E_1)-x'(E_2)\right)^2+\left(y'(E_1)-y'(E_2)\right)^2+\left(z'(E_1)-z'(E_2)\right)^2\right)
\end{aligned}
\tag{8}
$$

Es fácil comprobar que la ecuación (8) vale también bajo las transformaciones de los tipos (I) y (II). Por lo tanto, la cantidad expresada a ambos lados de la ecuación (8) permanece invariante bajo todas las transformaciones del grupo de Poincaré. De acuerdo con las ideas de Klein, este grupo determina entonces en $\mathbb{R}^4$ una geometría que llamaremos *minkowskiana*.[19] De esto depende decisivamente la noción de espaciotiempo que Minkowski aportó a la teoría de Einstein.

Para verlo es preciso entender bien qué es un sistema de coordenadas como los que hemos estado considerando. Cada sistema de estos asigna de modo exclusivo un cierto cuádruplo de números reales $\langle t(E),x(E),y(E),z(E)\rangle$ a un evento físico individual E, que ocupa un punto y dura un instante. Si suponemos con Minkowski que en el acontecer físico no hay vacíos,[20] cada uno de los sistemas considerados constituye en efecto una ☆aplicación biyectiva☆ $f: \mathcal{M} \to \mathbb{R}^4$, donde $\mathcal{M}$ —por 'mundo' o por 'Minkowski'— es el conjunto de los eventos. Digamos que una función f de este tipo es una *carta global* de $\mathcal{M}$ (f sería una carta *local* si estuviese definida solo en una *parte* de $\mathcal{M}$ y la aplicase sobre un *abierto* de $\mathbb{R}^4$). Si f es una carta global de $\mathcal{M}$, evidentemente existe la función inversa $f^{-1}: \mathbb{R}^4 \to \mathcal{M}$ que asigna a cada cuádruplo de números reales el evento al que f asigna ese cuádruplo. Una carta global f determina cuatro funciones $f^k: \mathcal{M} \to \mathbb{R}$, con valores reales $(0 \leq k \leq 3)$; el valor $f^k(E)$ de f^k en $E \in \mathcal{M}$ es la k-ésima coordenada de E en el sistema bajo consideración, por lo cual las funciones f^k también se llaman *coordenadas*. De ahora en adelante usaré las letras x, y,… para designar cartas —globales o locales— de $\mathcal{M}$, y distinguiré con superíndices numéricos las cuatro coordenadas correspondientes, poniendo x^0, y^0,… en vez de t, t',… ; x^1, y^1,… en vez de x, x',…; x^2,

y^2,… en vez de y, y',…, y x^3, y^3,… en vez de z, z',…[21] Estipularemos además que cualquier carta x que se mencione en lo que resta de la §1.3 cumplirá las siguientes condiciones: x^0 representa un tiempo de Einstein adaptado a un marco inercial de referencia $\mathcal{R}$, medido en segundos, y creciente en la dirección del futuro, mientras que x^1, x^2 y x^3 son coordenadas cartesianas orientadas por la derecha, que miden distancias en metros y están asociadas al mismo marco $\mathcal{R}$. Llamaré *carta estándar de Lorentz adaptada al marco inercial* $\mathcal{R}$ a cualquier carta que reúna estos requisitos. Una breve reflexión bastará para comprender que, si x e y son dos cartas estándar adaptadas, respectivamente, a los marcos inerciales $\mathcal{R}$ y $\mathcal{R}'$ (que bien podrían ser el mismo), la función compuesta $x \circ y^{-1}$ es una transformación de coordenadas que puede representarse mediante uno de los sistemas de ecuaciones (3), (4) o (7), o se compone de transformaciones que los satisfacen: en otras palabras, $x \circ y^{-1}$ es una transformación de coordenadas del grupo de Poincaré, y otro tanto vale para su inversa $y \circ x^{-1}$.[22] Llamemos $\mathcal{P}_{\mathbb{R}}$ a la realización del grupo (conectado) de Poincaré cuyos elementos son todas estas transformaciones. Sustituyendo cada función $x \circ y^{-1}$ —una permutación de $\mathbb{R}^4$— por la función $x^{-1} \circ y$, que es una permutación de $\mathcal{M}$, se obtiene otra realización del grupo (conectado) de Poincaré, isomorfa a la anterior, que llamaremos $\mathcal{P}_{\mathcal{M}}$. De acuerdo con las ideas de Klein, $\mathcal{P}_{\mathcal{M}}$ determina una geometría minkowskiana en el conjunto $\mathcal{M}$ de los eventos. Investido con esta estructura geométrica, $\mathcal{M}$ es lo que Minkowski llamó *mundo* (*Welt*) y hoy se llama *espaciotiempo de Minkowski*.

La geometría minkowskiana tiene una profunda afinidad con la euclídea, que percibiremos mejor en la §1.4; sin embargo, a primera vista, son muy diferentes. El invariante desplegado a ambos lados de la ecuación (8) asocia a cada par de eventos E_1 y E_2 una cantidad que denotaré con $\sigma^2(E_1,E_2)$ y que es patentemente la contraparte, en esta geometría, del cuadrado $\delta^2(P_1,P_2)$ de la distancia euclídea entre dos puntos P_1 y P_2. Pero la diferencia salta a la vista: mientras que $\delta^2(P_1,P_2) \geq 0$ y solo es igual a 0 si $P_1 \neq P_2$, de modo que no hay inconveniente en definir la distancia misma entre P_1 y P_2 como la raíz cuadrada *positiva* de $\delta^2(P_1,P_2)$, el invariante $\sigma^2(E_1,E_2)$ puede ser un número real *positivo*, *negativo* o *igual a cero*, aunque $E_1 \neq E_2$. Diré que la separa-

ción espaciotemporal entre dos eventos distintos E_1 y E_2 es *temporaloide* si $\sigma^2(E_1,E_2) > 0$, *espacialoide* si $\sigma^2(E_1,E_2) < 0$ y *nula* si $\sigma^2(E_1,E_2) = 0$.[23] La clasificación es inmediatamente aplicable a las curvas en el espaciotiempo, que llamaremos *temporaloides*, *espacialoides* o *nulas* si y solo si la separación entre cualesquiera dos de su puntos cae en la categoría homónima (hay, claro, curvas mixtas que no satisfacen esta condición). La analogía entre σ^2 y δ^2 puede acentuarse definiendo la "distancia minkowskiana" $\sigma(E_1,E_2)$ como la raíz cuadrada positiva de $|\sigma^2(E_1,E_2)|$ (una estipulación que también se deja aplicar, *mutatis mutandis*, sin inconveniente a la distancia euclídea). Pero la "distancia" así definida solo permite comparaciones significativas entre pares de eventos cuya separación es temporaloide, o alternativamente entre pares cuya separación es espacialoide, y no hace la menor diferencia entre los diversos pares cuya separación es nula. Ello no obstante, este concepto de distancia no carece de valor. Puede usárselo por ejemplo para definir la longitud de una curva temporaloide por analogía con la longitud de una curva cualquiera en el espacio euclidiano: si C es una curva temporaloide finita su longitud estaría dada —en términos de una carta estándar de Lorentz x— por la integral de línea

$$\int_C \sqrt{(dx^0)^2 - c^{-2}\left((dx^1)^2 + (dx^2)^2 + (dx^3)^2\right)} = \int_C \mathrm{d}\sigma \qquad (9)$$

Esta definición permite considerar a la curva C como análoga a una recta euclídea, en caso que $\int_C \mathrm{d}\sigma > \int_{C'} \mathrm{d}\sigma$ para cualquier otra curva vecina C' que una los extremos de C.[24] En tal caso, la integral (9) es estacionaria ($\delta\int_C \mathrm{d}\sigma = 0$) y decimos que C es una *geodésica* temporaloide. En general, si C es temporaloide, suele escribirse $\mathrm{d}\tau$ en vez de $\mathrm{d}\sigma$; la cantidad $\int_C \mathrm{d}\tau = \tau$ se llama el *tiempo propio* a lo largo de C.

Relativamente a un evento fijo cualquiera E, la función σ^2 divide el espaciotiempo en tres partes:

(i) el conjunto $\{X: \sigma^2(E,X) > 0\}$ de los eventos X cuya separación de E es temporaloide;

(ii) el conjunto $\{X: \sigma^2(E,X) < 0\}$ de los eventos X cuya separación de E es espacialoide;
(iii) el conjunto $\{X: \sigma^2(E,X) = 0\}$.

El conjunto (iii) es el llamado "cono de luz" de *E*, formado por dos hipersuperficies —dos ☆variedades tridimensionales☆— conectadas por el punto *E*, que hace de vértice. Los eventos de (i) temporalmente anteriores y posteriores a *E* forman, respectivamente, el *pasado cronológico* y el *futuro cronológico* de *E*. La unión del pasado cronológico de *E* y aquellos eventos de su cono de luz que pueden coincidir con la emisión de una señal luminosa recibida en *E* constituye el *pasado causal* de *E*. La unión del futuro cronológico de *E* y aquellos eventos de su cono de luz que pueden coincidir con la recepción de una señal luminosa emitida en *E* constituye el *futuro causal* de *E*. Algunos filósofos gustan decir que los eventos del conjunto (ii) son "topológicamente simultáneos con *E*", una denominación que puede aceptarse siempre que se tenga muy claro que esta relación de simultaneidad no es transitiva ni reflexiva: de $\sigma^2(E,X) < 0$ y $\sigma^2(X,Y) < 0$ no se infiere que $\sigma^2(E,Y) < 0$ y, por cierto, $\sigma^2(E,E) \not< 0$.

El nuevo punto de vista favorecido por Minkowski confiere un significado geométrico al principio de inercia. Si *p* es una partícula masiva sin volumen, la sucesión de eventos que forman la historia de su vida es el camino de una curva temporaloide, la *cosmolínea* de *p*.[25] Si *p* se mueve libremente, conforme al principio de inercia, su cosmolínea es una geodésica temporaloide.[26] Esta constatación permite caracterizar a cada marco inercial de referencia $\mathcal{R}$ como una congruencia☆ de tales geodésicas. A la luz de esta interpretación geométrica salta a la vista por qué los marcos inerciales y el tiempo de Einstein adaptado a cada uno de ellos gozan de una posición privilegiada en la *relatividad especial* o, dicho en otras palabras, en una física basada en la afirmación conjunta del principio de relatividad y el principio de la constancia de la velocidad de la luz.

1.4 Espaciotiempo y gravedad

La geometría minkowskiana hizo posible formular de un modo simple y elegante la electrodinámica clásica (Minkowski 1908) y la mecánica relativista (von Laue 1911*a*,*b*,*c*), pero en un principio Einstein no la miró con simpatía, refiriéndose a ella como un mero "formalismo", vale decir, como un artificio matemático capaz, como tantos otros, de facilitarle al físico sus cómputos pero no de enseñarle algo sobre la realidad que investiga.[27] Sin embargo, el espaciotiempo y su geometría son el cimiento y la sustancia de la teoría de la gravitación que llamamos *relatividad general*, que Einstein publicó a fines de noviembre de 1915.

Einstein inició la búsqueda de una teoría de la gravitación para sustituir a la de Newton en 1907, esto es, en el mismo año en que Minkowski dio a conocer por primera vez sus ideas en la Universidad de Göttingen (Minkowski 1915). Pero no empezó a utilizarlas hasta 1912, después de leer la lúcida exposición de esas ideas por Sommerfeld (1910*a*,*b*).[28] Sus primeras reflexiones sobre la gravedad se inspiran en las que antes lo llevaron a superar el problema de la electrodinámica de los cuerpos en movimiento. En el caso de la electrodinámica le llamaba la atención que procesos cuya apariencia es indiscernible se concebían de muy distinto modo según el marco inercial a que estuvieran referidos. Da este ejemplo: si un imán se mueve con velocidad **v** en presencia de un conductor en reposo, se piensa que en torno al imán surge un campo eléctrico con un determinado contenido energético que genera una corriente en los lugares por donde pasa el conductor; pero si es el conductor el que se mueve con velocidad –**v** en presencia de un imán en reposo se piensa que no surge campo eléctrico alguno en torno al imán, y que en cambio se produce en el conductor una fuerza electromotriz a la que no corresponde ninguna energía pero que genera en él una corriente igual a la generada en el otro caso (Einstein 1905*r*, p. 891). Con su principio de relatividad, Einstein puso término a esta asimetría de la electrodinámica. En 1907, advirtió una asimetría análoga en la física de la gravedad. ¿Qué diferencia fenoménica puede observarse entre la caída de los cuerpos en el interior (sin ventanas) de un ascensor en reposo en

un campo gravitacional que los acelera uniformemente con aceleración **g** y su inexorable movimiento hacia el piso del mismo ascensor si éste se mueve en ausencia de campos gravitacionales con aceleración –**g**? (Véase la fig. 1 en la p. 70.) Según la teoría de la luz que prevaleció en el siglo XIX la diferencia está en que un rayo de luz emitido desde una pared del ascensor en dirección paralela al piso mantendría esta dirección en el primer caso, mientras que en el segundo se curvaría ligeramente. Pero esta pretendida diferencia no tiene ni el más mínimo respaldo experimental, igual que no lo tenía el pretendido distingo entre las velocidades de la luz con respecto a un cuerpo que reposa y a otro que se mueve en el éter. Einstein adoptó por eso el siguiente *principio de equivalencia* que respeta la igualdad de los fenómenos: si $\mathcal{R}$ es un marco de referencia que reposa en un campo gravitacional homogéneo y $\mathcal{R}'$ es un marco de referencia que se mueve con aceleración constante en línea recta respecto a los marcos inerciales, la descripción de un experimento de física referida a cualquiera de los dos no permite distinguirlo del otro. Este enunciado del principio se inspira en la versión publicada por Einstein (1907*j*, p. 454) y repetida por él varias veces en escritos de divulgación. Pero para el lector de hoy, acostumbrado a ver en la televisión cosmonautas en caída libre dentro o fuera de una nave espacial, es mucho más intuitiva y elocuente otra formulación del principio de equivalencia, basada en lo que Einstein llama en un manuscrito póstumo "la idea más feliz de mi vida" ("der glücklichste Gedanke meines Lebens")[29]: si $\mathcal{R}$ es un marco de referencia que cae libremente en un campo gravitacional homogéneo y $\mathcal{R}'$ es un marco de referencia inercial, ningún experimento de física referido a uno de ellos permite decidir de cuál se trata.

Del principio de equivalencia, Einstein concluye, con audacia y sagacidad características, que inercia y gravedad son lo mismo. Newton asombró a todos equiparando el movimiento de la luna y los planetas con la caída de una manzana; Einstein va más lejos y los asimila a los movimientos inerciales a que Newton justamente los oponía. Con ello resuelve de una plumada el misterio que enturbia la teoría de la gravitación de Newton: ¿por qué la fuerza que causa la caída acelerada de los cuerpos es proporcional a la inercia con que ellos resisten

a la aceleración? Que todos los cuerpos, no importa su índole, experimenten la misma aceleración de gravedad se explica en la teoría de Newton solo por esta fortuita coincidencia. Pero ¿se la podrá preservar en la física einsteiniana, donde la inercia es relativa al marco de referencia? Max Planck (1907) estaba seguro que no y enseñó a distinguir conceptualmente entre "masa inercial" y "masa gravitacional" aunque experimentalmente se confundieran.[30] Einstein, en cambio, iluminado por "la idea más feliz de su vida", hizo de la identidad entre ambas el punto de partida inconmovible de su investigación de la gravedad.

La mecánica celeste newtoniana era a la sazón el capítulo más firmemente establecido de la física, pero se fundaba en la acción instantánea a distancia de fuerzas centrales de atracción entre todos los cuerpos materiales y era por lo tanto incompatible con la relatividad de la simultaneidad adoptada por Einstein para reconciliar la electrodinámica con la experiencia. Poincaré (1906) esbozó una ley de gravedad invariante bajo las transformaciones de Lorentz, pero Einstein comprendió muy pronto que por esa ruta se tropezaba con dificultades insalvables. Entonces, con su creación de 1905 todavía fresca y sin aceptar, se puso a trabajar en una teoría de la gravitación que decididamente la deja atrás. Su larga búsqueda pasó por varias etapas. A la formulación del principio de equivalencia y su aplicación a la óptica (1907*j*, 1911*h*), siguieron dos trabajos, hoy casi olvidados, sobre campos gravitacionales estáticos (1912*c*,*d*) y la "teoría generalizada de la relatividad" de Einstein y Grossmann (1913), a la que Einstein adhiere durante dos años pero que luego desecha en el frenesí de noviembre de 1915, cuando somete a la Academia Prusiana cuatro memorias, a razón de una por semana. En la primera de ellas propone una teoría de la gravitación (1915*f*), otra teoría diferente en la segunda (1915*g*), en la tercera demuestra que la segunda teoría predice el movimiento del planeta Mercurio mejor que la teoría de Newton (1915*h*), y solo en la última publica las ecuaciones del campo gravitacional propias de la relatividad general (1915*i*). Esta queda avalada también por los cálculos de la tercera semana porque casualmente coincide con la teoría de la segunda en el vacío interestelar donde circula el planeta. Aunque el feliz desenlace no estuvo asegurado hasta esa semana final, hoy

que ya lo conocemos podemos decir, sin temor a equivocarnos, que el paso crucial lo había dado Einstein poco antes de iniciar la colaboración con Grossmann, cuando comprendió que los fenómenos gravitacionales debían explicarse por la estructura geométrica del espaciotiempo. Según cuenta él mismo,[31] "la idea decisiva" de la analogía entre el problema que tenía entre manos y la teoría de las superficies de Gauss la tuvo al retornar a Zürich en el verano de 1912, como catedrático del Politécnico donde había hecho sus estudios. Su amigo Grossmann, que enseñaba matemáticas en la misma institución, le informó entonces que Riemann (1854) había generalizado la teoría de Gauss, extendiéndola a espacios de n dimensiones, y que para aplicarla a su problema Einstein tendría que valerse del cálculo diferencial absoluto —luego llamado cálculo tensorial— de Ricci y Levi-Civita (1901). Para entender el fundamento y la importancia de la analogía en cuestión debemos repasar rápidamente los logros de Gauss y considerar las dificultades que Einstein afrontaba inmediatamente antes de descubrirla.

Gauss (1828) estudia las superficies lisas discernibles en el espacio euclidiano de tres dimensiones. Introduce un método apropiado para investigar su "geometría intrínseca" sin atender al modo como la superficie está incrustada en el espacio. Piénsese, por ejemplo, en figuras dibujadas en una hoja de papel; la longitud de sus líneas y el tamaño de sus ángulos no cambia un ápice si la hoja se enrolla formando un cono o un cilindro. Si x^1 y x^2 son coordenadas cartesianas definidas sobre la hoja, la longitud $s(C)$ de cualquier curva C trazada en ella está dada por

$$s(C) = \int_C ds = \int_C \sqrt{(dx^1)^2 + (dx^2)^2} \tag{10}$$

La integral de la derecha puede escribirse de otro modo, utilizando el símbolo de Kronecker δ_{ij}, que designa el 1 si $i = j$ y el 0 si $i \neq j$:

$$s(C) = \int_C ds = \int_C \sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} \delta_{ij} dx^i dx^j} \tag{11}$$

Aunque no facilita la lectura de la ecuación (10), esta notación ayu-

dará a entender lo que sigue. Si x^1 y x^2 son coordenadas cartesianas, las líneas x^1 = const. y x^2 = const. —cualesquiera que sean las cantidades constantes que designamos con "const."— son rectas que se cortan ortogonalmente. Reemplacemos ahora las coordenadas cartesianas x^1 y x^2 con las coordenadas curvilíneas u^1 y u^2. Entonces las líneas u^1 = const. y u^2 = const. son curvas que forman ángulos variables en sus puntos de intersección. La longitud de la curva *C* puede sin duda expresarse en términos de las nuevas coordenadas, pero tenemos que reemplazar en la última integral los factores constantes δ_{ij} por factores que varían con la posición y, por lo tanto, se conciben como funciones de las coordenadas. Designándolos, como es habitual, con g_{ij} tenemos que

$$s(C) = \int_C ds = \int_C \sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} g_{ij} du^i du^j} \qquad (12)$$

El ☆elemento de línea☆ *ds* es por cierto igual aquí que en la ecuación (11), aunque su expresión analítica en el integrando de la derecha es distinta porque está referido a coordenadas de otro tipo. Consideremos ahora una superficie lisa cualquiera, por ejemplo, una esfera, o la superficie de un huevo o de una estatua de Henry Moore. En las superficies aducidas como ejemplos solo es posible definir coordenadas curvilíneas. Pero la longitud *s*(*C*) de una curva *C* trazada sobre cualquiera de ellas puede siempre representarse mediante la integral (12). En particular, si *C* es lo que se llama una *geodésica*, esto es, si las líneas adyacentes a *C* y trazadas entre sus dos extremos sobre la misma superficie son todas más largas o todas más cortas que *C*, *C* tiene que ser una solución del sistema de ecuaciones diferenciales que expresa las condiciones necesarias y suficientes para que

$$\delta \int_C \sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} g_{ij} du^i du^j} = 0 \qquad (13)$$

Obviamente, el elemento de línea $ds = \sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} g_{ij} du^i du^j}$ no puede aquí ser igual al que figura en la ecuación (11). Sin embargo, es oportuno señalar que coincide con aquel en la aproximación lineal, y la discrepancia solo se manifiesta a partir de las cantidades de segun-

do orden. En un lenguaje más intuitivo: la geometría intrínseca de una superficie curva no es la geometría del plano euclidiano, pero se le acerca óptimamente en un entorno infinitesimal de cada punto.

Gauss enseñó además a evaluar la curvatura propia de cada superficie mediante una función dependiente de las g_{ij}. Esta curvatura gaussiana es constante y positiva sobre una esfera, pero varía de un punto a otro en los otros casos citados: en el huevo obviamente crece hacia los extremos y es típicamente mayor en uno que en el otro; en una estatua antropomorfa, es mayor en los codos que en las palmas de las manos. Sobre una hoja de papel es constante e igual a cero, aunque la hoja se enrolle como cilindro; es negativa, en cambio, sobre el asiento de una montura, donde se cortan curvas cóncavas y convexas. Las funciones g_{ij} son pues portadoras de la información que especifica la geometría intrínseca de cada superficie. Por ejemplo, la geometría del plano euclidiano —parcialmente realizada en nuestra hoja de papel— puede caracterizarse así: en él y solo en él es posible definir coordenadas cartesianas x^1 y x^2 tales que las g_{ij} expresadas como función de ellas satisfacen la relación $g_{ij} = \delta_{ij}$.

Para hallar un parecido entre este tratamiento de las superficies y su propio problema con la gravedad Einstein tenía que adoptar el punto de vista geométrico de Minkowski como algo más que un mero formalismo, y concebir el espaciotiempo como la cancha o palestra en que se juega el acontecer cósmico. No sabemos exactamente cuando esto ocurrió, pero hay indicios de que fue muy poco antes del regreso de Einstein a Zürich. Hay un manuscrito póstumo suyo sobre la relatividad especial que empezó a escribir en 1912 para el *Handbuch der Radiologie* editado por Erich Marx (lo mencioné ya en la nota 27). A partir de la página 47 usó papel de fabricación suiza. En la página 46 califica las aportaciones de Minkowski como "conceptos auxiliares que simplifican extraordinariamente el sistema de la teoría de la relatividad",[32] una expresión que todavía deja traslucir cierta reserva. Pero en la página 45 leemos esto:

> Con Minkowski llamaremos "cosmopunto" (*"Weltpunkt"*) al punto cuadridimensional del espacio cuadridimensional (punto espaciotemporal) que corresponde a un determinado evento puntual y que está caracterizado, relativamente a un sistema de referencia admisible, por

coordenadas determinadas *x*, *y*, *z*, *u*.[33] *Este es propiamente el elemento de la descripción matemática de la naturaleza*. Si se trata, por ejemplo, de la descripción del movimiento de un punto material, este queda determinado si *x*, *y*, *z* están determinadas como funciones de *u*. Ello define una línea en el espacio cuadridimensional, que debemos llamar "cosmolínea" (*"Weltlinie"*) del punto material. La figura de esta línea no cambia si reemplazamos el sistema de referencia por otro igualmente admisible; solo varía la orientación de la línea entera relativamente al sistema de coordenadas.

(Einstein, CP, vol. 4, p. 68; énfasis mío)

En consonancia con este enfoque, en un suplemento añadido durante la corrección de pruebas a su artículo "Sobre la teoría del campo gravitacional estático" (1912*d*), Einstein escribe así la ecuación de la cosmolínea de un punto material que se mueve en un campo gravitacional estático, en ausencia de fuerzas externas:

$$\delta\left\{\int\sqrt{c^2dt^2-dx^2-dy^2-dz^2}\right\}=0 \qquad (14)$$

Si la velocidad de la luz *c* es constante y las coordenadas *t*, *x*, *y*, *z* son como las que usamos en la ecuación (9), esta es la ecuación de una geodésica temporaloide en el espaciotiempo de Minkowski, indistinguible de la cosmolínea de una partícula que se mueve libremente conforme al principio de inercia. Pero la teoría del campo gravitacional estático avanzada por Einstein supone que *c* depende del potencial gravitacional y solo es constante si este es constante. En el caso general es preferible hacer explícito que *c* varía con la posición, reescribiendo la ecuación (14) así:

$$\delta\left\{\int_C\sqrt{(c(x,y,z))^2dt^2-dx^2-dy^2-dz^2}\right\}=0 \qquad (14')$$

(donde he llamado *C* a la cosmolínea sobre la cual se toma la integral, para referirme a ella más tarde).

Esta variación de la velocidad de la luz con el potencial gravitacional la introdujo Einstein (1911*h*) como una consecuencia de su principio de equivalencia.[34] Max Abraham la incorporó en la nueva teoría de la gravedad que publicó en 1912 en varias entregas, a partir del mes de

enero (Abraham 1912*a*,*b*,*c*,*d*,*e*) y que Einstein criticó severamente, primero en cartas a Abraham y luego en un artículo entregado a *Annalen der Physik* el 26 de febrero (1912*c*).[35] En su presentación Abraham utiliza el formalismo espaciotemporal de Minkowski aunque su teoría no es invariante bajo las transformaciones de Lorentz ni siquiera en un entorno infinitesimal de cada evento puntual. Cuando Einstein le reprochó esta inconsecuencia, Abraham (1912*f*) se allanó a corregirla, admitiendo no sin cierta petulancia que su teoría prescindía de la relatividad einsteiniana, que el propio Einstein había ya desechado al abandonar el principio de la constancia de la velocidad de la luz.

En su respuesta a Abraham, Einstein (1912*h*) distingue con toda razón entre este principio y el de relatividad propiamente tal, que formula aquí expresamente como un principio de invariancia de las leyes naturales referidas a cualquier sistema de coordenadas de cierto tipo.[36] "Para dudar de la validez universal del principio de relatividad no tenemos […] ni el más mínimo fundamento. En cambio, soy de opinión de que el principio de la constancia de la velocidad de la luz solo se deja sostener si uno se limita a regiones espaciotemporales de potencial gravitacional constante" (p. 1062). Einstein admite, con todo, que parece imposible insertar consistentemente las fuerzas de gravedad "en el esquema de la teoría de la relatividad de hoy" (*das Schema der heutigen Relativitätstherie*—p. 1063). En el interesantísimo párrafo siguiente (el penúltimo de su respuesta), añade que "la teoría de la relatividad de hoy siempre conservará su valor, en mi opinión, como la teoría más simple *en el importante caso límite* del acontecer espaciotemporal con potencial gravitacional constante" (ibid., cursiva mía). Como tarea para "el futuro inmediato", Einstein propone "crear un esquema relativista (*ein relativitätstheoretisches Schema*) en que se exprese la equivalencia entre masa inerte y masa pesada". Sus trabajos sobre el campo gravitacional estático constituyen solo un primer aporte, "muy modesto", al logro de esta meta. En ellos ha procurado asegurar la equivalencia de las masas asimilando el campo gravitacional estático a un marco de referencia acelerado. Einstein reconoce que solo ha podido llevar a cabo esta asimilación sin contradicciones en espacios infinitesimales, y que no sabe explicarse esta limitación. Pero "no ve aquí ninguna razón para descartar dicho principio de equiva-

lencia también para lo infinitesimal: nadie puede negar que este principio es una extrapolación natural de uno de los principios empíricos más universales de la física." Sigue una observación que destaco para poder luego referirme a ella más fácilmente:

> Por otra parte, este principio de equivalencia nos abre la interesante perspectiva de que las ecuaciones de una teoría de la relatividad que también comprendiera la gravitación pudiesen ser invariantes también respecto de las transformaciones de aceleración (y rotación).
>
> (Einstein, 1912*h*, p. 1063)

Enseguida agrega: "A la luz del caso sumamente especial considerado hasta aquí de la gravitación de masas en reposo se puede ver ya que *las coordenadas espacio-temporales tienen que perder su simple significado físico*" (ibid., pp. 1063–64, cursiva mía). En la Gibson Lecture dictada en Glasgow en 1933, Einstein dirá que una de los peores obstáculos que tuvo que vencer cuando buscaba una teoría de la gravitación fue su propia resistencia a aceptar la conclusión que he puesto en cursiva.[37] Ahora bien, si la ecuación (14) se mira desde este punto de vista, la analogía con la teoría gaussiana de las superficies empieza a parecer irresistible. Reemplacemos en ella las letras t, x, y, z con x^0, x^1, x^2, x^3, como hicimos para hablar del espaciotiempo de Minkowski. Si ponemos $g_{00} = c(x^1,x^2,x^3)$, $g_{kk} = -1$ si $k > 0$, y $g_{kh} = 0$ si $k \neq h$, la ecuación (14′) pasa a escribirse así:

$$\delta\int_C \sqrt{\sum_{i=0}^{3}\sum_{j=0}^{3} g_{ij}dx^i dx^j} = 0 \qquad (15)$$

A primera vista, esta ecuación no se distingue de la ecuación (13) de una geodésica sobre una superficie curva más que por el número de dimensiones en juego.[38] En principio, la ecuación (15) podría tal vez representar una geodésica del espaciotiempo de Minkowski, referida a coordenadas curvilíneas; pero esta alternativa está excluida si el potencial gravitacional no es constante. Sin embargo, la analogía gaussiana induce a concebir (15) como la ecuación de una geodésica de un espaciotiempo con otra geometría, que se relacione con la de Minkowski como la geometría de una superficie curva se relaciona

con la del plano euclidiano. Si, aplicando la "extrapolación natural" favorecida por Einstein, suponemos que el principio de equivalencia vale "para lo infinitesimal", el integrando $\sqrt{\sum_{i=0}^{3}\sum_{j=0}^{3} g_{ij}dx^i dx^j}$ de la ecuación (15) se aproximará linealmente en un pequeño entorno de cada evento al integrando dσ de la ecuación (9), esto es, al "elemento de línea" de la geometría minkowskiana, y la cosmolínea de una partícula en caída libre en un campo gravitacional cualquiera no se distinguirá localmente de lo que sería la cosmolínea de una partícula en movimiento inercial. (Considérese este ejemplo familiar: un cosmonauta en órbita deja caer un martillo; este se comporta a ojos vistas, de momento en momento, como un cuerpo libre de toda influencia de fuerzas externas, aunque cae en el campo gravitacional de la Tierra, que no es homogéneo).

Falta mencionar el aspecto más inesperado y fecundo de la analogía gaussiana. Extendiendo audazmente el principio de equivalencia a los campos gravitacionales inhomogéneos, Einstein concibe la inercia como un caso límite de la gravedad y la ley de la caída libre como una generalización de la ley de inercia. Según esta ley, la cosmolínea de una partícula en movimiento inercial es una geodésica temporaloide en el espaciotiempo. Einstein postula que otro tanto vale —*mutatis mutandis*— para la cosmolínea de una partícula en caída libre.[39] La diferencia —los *mutanda*— reside solo en la geometría aplicable en cada caso y se expresa en las g_{ij}, que en el caso de la inercia pueden siempre reducirse, mediante una juiciosa elección de coordenadas, a la colección de constantes desplegada en la matriz siguiente

$$\begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad (16)$$

pero que en el caso general son funciones de las coordenadas y varían con la posición espaciotemporal. Como la cosmolínea geodésica de una partícula que cae libremente está enteramente determinada por las

g_{ij}, estas funciones, portadoras de la información que caracteriza la geometría del espaciotiempo, especifican a la vez el campo gravitacional, que guía la trayectoria espaciotemporal de cada cuerpo expuesto a su influencia. Así, en la futura teoría de la gravedad, las g_{ij} jugarán todas juntas el papel del potencial gravitacional en la teoría clásica. Tal como esta vincula, mediante la ecuación de Poisson, las variaciones del gradiente del potencial con un campo escalar que representa la materia presente en cada lugar, la teoría contemplada por Einstein descansará en un sistema de ecuaciones del campo gravitacional que vincula las g_{ij} y sus derivadas primeras y segundas a un objeto matemático representativo de la distribución de la energía no gravitacional.

Las obras de Riemann y de Ricci y Levi-Civita que Einstein estudia por consejo de Grossmann no podían sino confirmarlo en su visión, inspirada en Gauss, de una teoría geométrica —o, mejor dicho, geometrodinámica— de la gravedad. La historia de sus afanes y extravíos —desde agosto de 1912, cuando ya todo lo esencial estaba claro para él, y fines de noviembre de 1915, cuando al fin dará con las ecuaciones de campo que buscaba— ha sido bien narrada varias veces y no viene al caso repasarla aquí.[40] La teoría de la gravitación que finalmente propone significa el triunfo del punto de vista espaciotemporal de Minkowski sobre el enfoque inicial de Einstein, basado en la comparación entre marcos de referencia intersustituibles. Cuesta por eso entender la tenacidad del nombre "teoría general de la relatividad" que Einstein le impuso. Ya a propósito de la "teoría especial" había dicho Minkowski que le parecía "muy insípido" (*sehr matt*) llamar

> *Principio de Relatividad* (*Relativitätspostulat*) a la exigencia de invariancia bajo el grupo [de Lorentz]. Como el sentido del principio es que con los fenómenos solo se nos da el cosmos de cuatro dimensiones en el espacio y tiempo, mientras que su proyección en el espacio y en el tiempo puede siempre llevarse a cabo con cierta libertad, preferiría darle a este aserto el nombre de *Principio del mundo absoluto* (*Postulat der absoluten Welt*).
>
> (Minkowski, 1909, p. 107)

La “teoría general” abarca ciertamente a la teoría especial de la relatividad como un caso límite, pero solo porque las g_{ij} desplegadas en la matriz (17), que caracterizan la geometría minkowskiana, constituyen una solución particular de las ecuaciones de Einstein (1915*i*) —aplicable en el caso de un espaciotiempo vacío— y esta solución es aproximada linealmente por cualquier otra en un entorno de cada evento puntual. Es pues en cuanto teoría del espaciotiempo que la teoría general merece este calificativo, y no en cuanto pretendida generalización del principio de relatividad. La observación de Einstein (1912*h*) que destaqué arriba manifestaba la esperanza de extender la relatividad a los marcos de referencia en movimiento acelerado. La idea de que esta esperanza se ha cumplido con creces en cuanto las ecuaciones de Einstein (1915*i*) satisfacen el requisito de covariancia general es fruto de una confusión, como el mismo Einstein (1918*e*) reconoció en su respuesta a Kretschmann (1917), que se lo hizo ver. La covariancia general, esto es, la propiedad que tienen esas ecuaciones de preservar su validez bajo transformaciones arbitrarias de coordenadas, es una exigencia que naturalmente deben cumplir las leyes de la física cuando se las refiere a sistemas de coordenadas que, en principio, carecen de significado métrico y se usan solo para individualizar los eventos y expresar las relaciones de vecindad entre ellos mediante asignaciones de números. Esta exigencia se cumple, por lo demás, automáticamente cuando las cantidades físicas se representan mediante “objetos geométricos”[41] —como los tensores del cálculo diferencial de Ricci y Levi-Civita (1901) que debe justamente a esta circunstancia su epíteto de *absoluto*— y no se opone en modo alguno a la construcción de un “objeto geométrico” que corresponda a una rotación absoluta, como lo requiere, por ejemplo, la solución de Gödel (1949) a las ecuaciones de Einstein (1915*i*).

2

El orden del tiempo

Este capítulo se refiere a las innovaciones que Einstein introdujo con sus teorías de la relatividad en nuestra concepción del orden del tiempo.[1]

2.1 Simultaneidad y sucesión en la relatividad especial

Mucho antes de que aprendiéramos a coordinar nuestros quehaceres diarios mediante agendas y relojes, los antiguos astrónomos le asignaron coordenadas temporales a las posiciones observadas y previstas de los "astros errantes" (*planetai*, inclusive el sol y la luna), correlacionándolas con la rotación del firmamento alrededor de los polos. Mientras se creyó que el firmamento era una esfera inmutable que contenía a todo el universo y que la luz llenaba instantáneamente cualquier medio transparente expuesto a ella, cabía considerar cualquier etapa de la revolución de las "estrellas fijas" como simultánea con todos los sucesos que uno percibía a la misma vez que ella. Pero ello se hizo imposible después que Copérnico, Newton, Rømer y otros establecieron que (i) las "estrellas fijas" no están clavadas en una esfera sino que están libremente repartidas en el espacio sin límites y probablemente se mueven unas respecto de otras; (ii) la aparente rotación del cielo es en realidad una rotación de la tierra, que no puede ser perfectamente uniforme; (iii) la luz se propaga con una velocidad finita. No obstante, todos los físicos y astrónomos aceptaban tácitamente que "cada momento de duración se difunde indiviso por todo el espacio universal" (como dice Newton en un manuscrito publicado póstumamente en 1962).[2] Hasta donde sé, el primero que vio aquí una dificultad fue James Thomson, el hermano mayor de Lord Kelvin. Según él,

> Tenemos muy buenos medios para conocer en algunos casos e imaginarnos en otros la distancia entre los puntos del espacio ocupados simultáneamente por los centros de dos pelotas; al menos, si nos conformamos con dejar a un lado la dificultad concerniente a la imperfección de nuestros medios para establecer o especificar, o para idealizar claramente, la simultaneidad en lugares distantes entre sí. Usamos de ordinario para ello señales acústicas, luminosas, eléctricas, o enviadas por conexiones de alambre o barras, o por varios otros medios. El tiempo empleado en la transmisión de la señal implica una imperfección en las facultades humanas para establecer la simultaneidad de sucesos que ocurren en lugares distantes entre sí. Parece, sin embargo, que probablemente no haya ninguna dificultad en idealizar o imaginar la existencia de la simultaneidad.
>
> (James Thomson 1884, p. 569)

Muy poco después, la dificultad que James Thomson "se conformaba con dejar a un lado" se tornó aguda gracias a un instrumento óptico de gran precisión inventado por A. A. Michelson (1881). El fracaso de Michelson en sus intentos para detectar con su aparato algún indicio del "movimiento relativo de la tierra y el éter luminífero" tuvo un efecto catastrófico sobre la teoría vigente de la propagación de la luz. En su afán de rescatar esta teoría —y con ella la electrodinámica clásica de Maxwell, de la que parecía ser un ingrediente esencial— H. A. Lorentz no tuvo empacho en intervenir en el modo estándar de computar el tiempo. Sean t y $\mathbf{r}(t)$, respectivamente, la coordenada temporal estándar y el vector de posición asociados al marco referencial del éter. Sea c la velocidad de la luz en el vacío, referida a $\mathbf{r}$ y t. Consideremos un sistema que se mueve a través del éter con la velocidad constante $\mathbf{v}$, y supongamos que el vector de posición $\mathbf{r}'$ asociado a este sistema está dado por la transformación estándar llamada de Galileo:

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t \tag{1}$$

Lorentz (1895) introdujo una variable de "tiempo local" t', que satisface la ecuación

$$t' = t - c^{-2}(\mathbf{v} \cdot \mathbf{r}) \tag{2}$$

El *teorema de los estados correspondientes* de Lorentz puede entonces enunciarse en términos de las coordenadas con y sin acento:

> Si hay un sistema de cuerpos en reposo en el éter y una solución de las ecuaciones del campo electromagnético tal que el desplazamiento dieléctrico, la fuerza eléctrica y la fuerza magnética en ese marco sean ciertas funciones de $\mathbf{r}$ y t, hay una solución para el mismo sistema de cuerpos cuando se mueve relativamente al marco del éter con velocidad $\mathbf{v}$, tal que el desplazamiento dieléctrico, la fuerza eléctrica y la fuerza magnética en el marco inercial en que el sistema reposa son exactamente las mismas funciones de $\mathbf{r}'$ y t'.[3]

Lorentz nada dice sobre el significado físico del tiempo local y años más tarde observará que no es más que "una cantidad matemática auxiliar". Sin embargo, el teorema de los estados correspondientes no serviría a ningún propósito útil si el tiempo local no tuviera una relación bien definida con la marcha de los relojes utilizados en los laboratorios. Pero Lorentz no hizo el menor amago de explicar por qué tales relojes se ven afectados, y afectados del mismo modo, al transportarlos a través del éter.

La obra de Lorentz puede haber provocado ciertas reflexiones sobre la medición del tiempo publicadas por Henri Poincaré en 1898. Es oportuno recordar que, aunque inventivo y lleno de recursos, Lorentz, que nació en 1853, estaba enteramente aclimatado en lo que hoy llamamos física clásica; Poincaré, en cambio, aunque solo nueve meses menor que Lorentz, era esencialmente un matemático y, como tal, un pensador mucho más libre de lo que solían ser en ese entonces los cultivadores de una ciencia natural. En sus trece secciones, el ensayo de Poincaré aborda muchas cuestiones importantes. Me referiré solo a tres.

Hay ante todo un problema con la longitud de los intervalos de tiempo. ¿Qué nos autoriza a decir que tal proceso, que ocurrió ayer, y tal otro, que ocurrió el día en que Newton cumplió siete años, duraron ambos exactamente tres segundos? En tiempos de Poincaré, los astrónomos medían el tiempo por las oscilaciones de relojes de péndulo, cuyo período postulaban invariable, en primera aproximación; pero estaban de acuerdo en que lo afectan los cambios de la temperatura ambiente y de la presión atmosférica y quizás también

otras fuerzas, aún no detectadas. "De hecho, los mejores relojes tienen que ser corregidos de tiempo en tiempo, y las correcciones se hacen con ayuda de las observaciones astronómicas; uno se las arregla para que el reloj sideral dé la misma hora cada vez que la misma estrella cruza el meridiano. En otros términos, el día sideral, esto es, el período de rotación de la tierra, es la unidad constante del tiempo" (Poincaré 1898, III; en Poincaré VS, p. 43). Pero los astrónomos no se dan por contentos con esta definición. "Muchos piensan que las mareas actúan como un freno sobre nuestro globo" (ibid.). Invocan así las leyes de la dinámica, o, en otras palabras, suponen que "el tiempo debe definirse de tal modo que la ley de Newton y el principio de conservación de la energía se verifiquen" (1898, V; en VS, p. 46). Sin embargo, si uno adoptase un modo diferente de medir el tiempo, ello no cambiaría los fenómenos en que se basa la ley de Newton. "Solo el enunciado de la ley sería diferente, puesto que estaría traducido a otro lenguaje; evidentemente sería mucho menos simple. De modo que la definición implícitamente adoptada por los astrónomos puede resumirse así: El tiempo debe *definirse de tal modo que las ecuaciones de la mecánica sean tan simples como se pueda.* Dicho de otro modo: no hay una manera de medir el tiempo que sea más verdadera que otra; la que generalmente se adopta es solo la más cómoda" (ibid.; cursiva mía).

Luego está la cuestión del orden del tiempo. Según Poincaré hay una regla que es "la única que podemos seguir" en esta materia: "cuando un fenómeno se nos presenta como la causa de otro, lo consideramos como anterior" (1898, IX; en VS, pp. 49s.) Pero ¿qué hacer en un caso como el siguiente? En 1572 Tycho Brahe vio en el cielo la explosión de una nova; pero la luz de esa explosión tiene que haberse demorado por lo menos doscientos años en llegar a la tierra. Como ninguna señal conocida se transmite más rápidamente que la luz, esa nova no puede haber tenido ninguna influencia sobre el descubrimiento de América por los españoles en 1492. "Cuando digo que ese fenómeno es anterior a la formación de la imagen visual de la isla Española en la conciencia de Cristóbal Colón ¿qué es lo que quiero decir? Basta un poco de reflexión para comprender que esta

aseveración no tiene por sí misma ningún significado. Solo puede adquirirlo mediante una convención" (1898, VI; en VS, p. 47).

Por último, tenemos el problema de la simultaneidad entre sucesos distantes. Poincaré sabía, por cierto, que en su tiempo este se resolvía normalmente acarreando por toda la tierra relojes sincronizados, por ejemplo, en París. De este modo, dice, "el problema cualitativo de la simultaneidad se reduce al problema cuantitativo de la medición del tiempo" (1898, XII; en VS, p. 53). No insiste en las dificultades inherentes a esta, que ya examinó antes; ni menciona una complicación probable, a saber, que el ritmo de un reloj puede ser diversamente afectado según la velocidad con que se lo transporta de un sitio a otro. Pero se detiene a considerar una nueva regla para establecer la simultaneidad, completamente distinta de las conocidas hasta entonces, basada en el postulado —"aceptado por todo el mundo" (ibid.)— de la constancia de la velocidad de la luz en todas direcciones. "Este postulado no podrá jamás ser verificado directamente por la experiencia; podría ser contradicho por ella, si los resultados de las diversas mediciones no fuesen concordantes. Debemos alegrarnos de que tal contradicción no se produzca y que las pequeñas discordancias puedan explicarse fácilmente", ya que "es un postulado sin el cual no podría intentarse una medición de [la] velocidad [de la luz]" (1898, XII; en VS, p. 52).

En el § 1 del artículo en que introdujo lo que más tarde se llamará la teoría especial de la relatividad, Einstein (1905*r*) aborda sin rodeos la cuestión de la simultaneidad a distancia.[4] El propósito declarado de este artículo es mostrar que el principio newtoniano de la relatividad, según el cual "ninguna propiedad de los fenómenos corresponde al concepto de reposo absoluto", vale "no solo en la mecánica, sino también en la electrodinámica" (Einstein 1905*r*, p. 891).[5] Einstein quiere combinar este principio con el supuesto de que "la luz se propaga siempre en el vacío con una velocidad determinada c, que es independiente del estado de movimiento del cuerpo que la emite" (1905*r*, p. 892; reemplazo la letra V empleada por Einstein por la letra c, ahora estándar). Me referiré a estos dos supuestos de Einstein como el *primero* y el *segundo principio de la relatividad especial*. Einstein dice que su aseveración conjunta es suficiente para fundar

una electrodinámica simple y consistente de los cuerpos en movimiento, acorde con la teoría de Maxwell para los cuerpos en reposo. Para demostrar este aserto, Einstein tiene primero que probar que los dos principios de la relatividad especial no se contradicen mutuamente. Porque, si una señal luminosa se propaga con la velocidad $\mathbf{v}$ relativamente a un cuerpo que se desplaza respecto a nosotros con la velocidad constante $\mathbf{w} \neq 0$, ¿no debería esa misma señal propagarse, *relativamente a nosotros*, con la velocidad $\mathbf{v} + \mathbf{w}$, diferente de $\mathbf{v}$?

Para disipar estas nociones previas, Einstein emprende un examen crítico de la cinemática, esto es, de nuestro modo de describir el movimiento. Tal como se hacía corrientemente, Einstein refiere su descripción a un marco inercial de referencia que llamaré $\mathcal{R}$.[6] La posición de una partícula P que reposa en $\mathcal{R}$ puede determinarse relativamente a este marco "mediante varas rígidas, utilizando los métodos de la geometría euclídea, y expresarse en coordenadas cartesianas" (1905*r*, p. 892). Sin embargo, para describir el movimiento de una partícula tenemos que dar sus coordenadas de posición como función del tiempo. Esta descripción matemática solo tiene sentido si ya nos hemos formado una idea clara de lo que aquí significa 'tiempo'. Einstein da por descontado tácitamente que hay relojes naturales que marchan al unísono (esto es, de tal modo que siempre se mantiene la misma proporción entre sus períodos respectivos). Mediante un reloj natural situado en cualquier punto *A* de $\mathcal{R}$ se le pueden asignar coordenadas temporales a todos los eventos que ocurren en *A*. Se puede hacer otro tanto con los eventos que ocurren en otro punto *B,* mediante un reloj situado en *B*. Pero no es posible describir apropiadamente el movimiento de una partícula P desde *A* hasta *B* a menos que uno pueda comparar el tiempo en que P sale de *A* con el tiempo en que llega a *B*. Para ello no basta que los relojes en *A* y en *B* marchen al unísono; es preciso además que asignen las mismas coordenadas temporales a eventos simultáneos entre sí.

Einstein propone, medio en broma, la convención siguiente:[7] si estoy en reposo en el punto *A* de $\mathcal{R}$, le asignaré a cualquier evento *E* el tiempo que indica un reloj que reposa en *A* en el instante en que veo *E*. Einstein descarta inmediatamente esta convención, porque hace que la coordenada temporal de cualquier evento dependa del punto

desde el cual se lo observa. Tal convención sería por cierto muy incómoda. Pero merece además un reparo mayor que Einstein no menciona: un sistema de coordenadas en que la simultaneidad se establece de este modo no cumple el requisito mínimo que Einstein impone a $\mathcal{R}$ (citado en la nota 6, p. 212). En efecto, si se usa el tiempo así definido para determinar la velocidad de una partícula libre que viene hacia mí, dicha velocidad sufrirá un cambio discontinuo cuando la partícula pase a mi lado.[8]

Einstein no se molesta en mencionar siquiera el método utilizado a la sazón en la práctica para fijar el tiempo universal, al menos en la superficie de la tierra, mediante el transporte de relojes, y pasa directamente a describir el método —que hoy lleva su nombre— para sincronizar relojes distantes mediante señales de radar. Sean R_A y R_B dos relojes que reposan en $\mathcal{R}$, respectivamente, en los puntos A y B. Para sincronizar R_B con R_A, se envía una señal luminosa desde A hasta B cuando R_A indica el tiempo t_1. Si la señal, tras rebotar en B, llega de vuelta a A cuando R_A indica el tiempo t_3, el reloj R_B se ajustará de tal modo que a la llegada de la señal a B habría indicado el tiempo t_2, donde, *por definición*,

$$t_2 = t_1 + \tfrac{1}{2}(t_3 - t_1) \qquad (3)$$

Así, la señal se demora lo mismo en ir de A hasta B que en regresar desde B a A. Como B es un punto cualquiera de $\mathcal{R}$, las señales utilizadas para establecer una coordenada temporal de esta manera satisfacen el segundo principio de la relatividad especial *por definición*, siempre que sea posible ejecutar coherentemente el procedimiento. Como dije en el capítulo 1, llamo *tiempo de Einstein adaptado a* $\mathcal{R}$ a una coordenada temporal definida de este modo. Llamaré *carta de Lorentz adaptada al marco de referencia* $\mathcal{R}$ a un sistema de coordenadas (t,x,y,z), formado por un tiempo de Einstein t definido para $\mathcal{R}$ como acabo de explicar, y coordenadas cartesianas (x,y,z) tales que las ecuaciones $x = 0$, $y = 0$ y $z = 0$ definen tres planos mutuamente ortogonales que reposan en $\mathcal{R}$. Salvo que advierta otra cosa, supongo que las coordenadas cartesianas están siempre orientadas "por la derecha", en el sentido definido en la p. 25.[9]

El expediente a primera vista facilón de que se vale Einstein para asegurar la constancia de la velocidad de la luz preocupó a físicos y filósofos durante casi un siglo. Mientras que estos insistían casi todos en que solo la velocidad media de la luz en un viaje *de ida y regreso* puede considerarse como una cuestión de hecho, aquellos se afanaban en hallar un método para medir su velocidad en una sola dirección sin acordar primero un procedimiento para sincronizar los relojes a distancia. La cuestión fue resuelta por la Conferencia Internacional de Pesos y Medidas cuando decidió en 1983 que la unidad de longitud, el *metro*, es igual a 1/299.792.458 de la distancia que recorre la luz en el vacío en un segundo (la unidad internacional de tiempo). En virtud de esta decisión, y mientras ella esté vigente, no hay la más mínima posibilidad de propagar la luz en el vacío a una velocidad mayor o menor que 299.792.458 metros por segundo. Por lo tanto, los métodos experimentales propuestos para medir la velocidad unidireccional de la luz, si sirven para algo, sería solo para verificar que la distancia $\delta(A,B)$ desde un punto A hasta un punto B es igual a la distancia $\delta(B,A)$ desde B hasta A. Por otra parte, los filósofos convencionalistas pueden siempre refocilarse pensando que ninguna convención internacional ha restringido su libertad para desechar la definición estándar de la distancia como una función simétrica sobre pares de puntos y estipular que $\varepsilon\delta(A,B) = (1 - \varepsilon)\delta(B,A)$ para algún número real ε menor 1 y diferente de 0.5, que dependa de la dirección $\overrightarrow{AB}$.[10]

Curiosamente, mientras la cuestión de la convencionalidad de la velocidad unidireccional de la luz era todavía un problema vivo —esto es, mientras la definición del metro se basaba en la barra de platino iridiado que se conserva en Sèvres (1889–1963), o en la longitud de onda de la radiación emitida entre dos niveles de energía bien definidos del Criptón-86 (1963–1983)— los partícipes en el debate prestaron muy poca atención a las condiciones de coherencia impuestas por Einstein en su definición de simultaneidad (3). Requiere expresamente que dicha definición pueda aplicarse desde cualquier punto del marco inercial $\mathcal{R}$ de tal modo que (i) si un reloj situado en B se sincroniza con uno situado en A mediante una señal enviada desde A y otro reloj en A se sincroniza con el reloj en B mediante una

señal enviada desde *B*, los dos relojes en *A* quedan sincronizados entre sí, y (ii) si un reloj en *B* y otro en *C* se sincronizan con un reloj en *A* mediante una señal enviada desde *A* y otro reloj en *C* se sincroniza con el reloj en *B* mediante una señal enviada desde *B*, los dos relojes en *C* quedan sincronizados entre sí (1905*r*, p. 894). También, sin duda, supuso tácitamente que su definición de simultaneidad puede aplicarse en cualquier momento de tal modo que si un reloj en *B* se sincroniza con otro en *A* mediante una señal enviada desde *A* en un instante t, y luego otro reloj en *B* es sincronizado con el reloj en *A* mediante una señal enviada desde *A* en un instante posterior t', los dos relojes en *B* quedan sincronizados entre sí. (Esto presupone la premisa tácita de Einstein de que los relojes naturales marchan de modo uniforme y concordante, pero no se deduce de ella sin más). Además, para que el marco $\mathcal{R}$ satisfaga el requisito mínimo prescrito por Einstein (véase la nota 6, p. 212), las partículas libres que no reposen en $\mathcal{R}$ tienen que moverse con velocidad constante medida según el tiempo de Einstein; en otras palabras, si $\mathbf{r}$ es el vector de posición de una partícula libre en el espacio euclidiano determinado por $\mathcal{R}$ y t es el tiempo de Einstein, $\mathrm{d}^2\mathbf{r}/\mathrm{d}t^2 = 0$.[11] Todas estas son condiciones fácticas que se cumplen en la medida en que las predicciones de la relatividad especial, referidas a una carta de Lorentz, se verifican experimentalmente. Si las condiciones señaladas no se cumplieran, la definición de simultaneidad (3), aunque lógicamente posible, sería físicamente inútil (y los metrólogos no habrían adoptado la definición actual del metro). Aunque las convenciones son libres, su idoneidad depende de la naturaleza de las cosas. Por ejemplo, la convención de servir la comida en platos no es idónea si el invitado es una cigüeña.

Evidentemente, es posible transformar las coordenadas de dos cartas de Lorentz adaptadas al mismo marco inercial ajustando las unidades de medida, transladando el origen de la coordenada temporal y de las coordenadas espaciales y rotando en torno a su vértice el triedro que determina las coordenadas espaciales. Todas estas transformaciones son en cierto modo triviales y las llamaré así. Consideremos, en cambio, la relación entre cartas de Lorentz adaptadas a dos marcos inerciales $\mathcal{R}$ y $\mathcal{R}'$, tales que $\mathcal{R}'$ se mueve relativamente a $\mathcal{R}$ con velocidad constante $\mathbf{v}$. Mediante transformaciones triviales es posible

reducir cualquier par de cartas de Lorentz adaptadas a $\mathcal{R}$ y $\mathcal{R}'$, respectivamente, a sendas cartas de Lorentz $X = (t,x,y,z)$ y $X' = (t',x',y',z')$ que utilizan las mismas unidades de medir y cumplen con las condiciones siguientes:

(i) ambas cartas asignan las coordenadas (0,0,0,0) al mismo evento;
(ii) en el instante $t = t' = 0$, los tres ejes de las coordenadas espaciales de X tienen la misma dirección y orientación que los ejes homónimos de X'; y
(iii) si O′ es el origen de las coordenadas espaciales de X' en $\mathcal{R}'$, tenemos que $\frac{dx(\mathrm{O}')}{dt} = |\mathbf{v}|$, y por lo tanto

$$\frac{d^2 y(\mathrm{O}')}{dt^2} = \frac{d^2 z(\mathrm{O}')}{dt^2} = 0$$

Entonces, la aseveración conjunta de los dos principios de la relatividad especial implica que las cartas X y X' están relacionadas por la siguiente *transformación cinemática de Lorentz*:

$$\begin{aligned} t' &= \frac{t - \left(vx/c^2\right)}{\sqrt{1 - \left(v/c\right)^2}} \\ x' &= \frac{x - vt}{\sqrt{1 - \left(v/c\right)^2}} \\ y' &= y \\ z' &= z \end{aligned} \tag{4}$$

donde c es la velocidad de la luz en el vacío y $v = |\mathbf{v}|$ es la velocidad de $\mathcal{R}'$ en $\mathcal{R}$.[12]

Ello tiene consecuencias inesperadas que a muchos les costó asimilar, porque les parecieron reñidas con su intuición del tiempo. Basta dar una rápida mirada a las ecuaciones (4) para comprender que dos eventos simultáneos en $\mathcal{R}$ pueden no ser simultáneos en $\mathcal{R}'$ (y, por

tanto, que dos eventos que se suceden en $\mathcal{R}$ pueden darse en $\mathcal{R}'$ con el orden de sucesión invertido).[13] Pero la verdad es que no tenemos ninguna intuición referente a la simultaneidad de eventos distantes. Más problemática parecía ser la siguiente consecuencia de las ecuaciones (4). Considérense dos eventos E y E^* que ocurren en el origen de las coordenadas espaciales de la carta X. El intervalo de tiempo entre ambos está dado en términos de X por $|t(E) - t(E^*)|$ y por $|t'(E) - t'(E^*)|$ en términos de X'. Estos dos números no son iguales, pues

$$\begin{aligned} |t'(E) - t'(E^*)| &= \frac{|t(E) - t(E^*) - (v/c)(x(E) - x(E^*))|}{\sqrt{1-(v/c)^2}} \\ &= \frac{|t(E) - t(E^*)|}{\sqrt{1-(v/c)^2}} > |t(E) - t(E^*)| \end{aligned} \tag{5}$$

(a menos que $v = 0$). Por tanto, un reloj que reposa en $\mathcal{R}'$, comparado con una fila de relojes sincronizados por el método estándar de Einstein en reposo en $\mathcal{R}$, a lo largo de los cuales pasa, parecerá marchar más lento que estos.[14] Se creyó descubrir aquí una contradicción, pues, aplicando la inversa de la transformacion (4), un reloj que reposa en $\mathcal{R}$, comparado con una fila de relojes sincronizados por el método estándar de Einstein en reposo en $\mathcal{R}'$, a lo largo de los cuales pasa, tendría que verse marchando más lento que estos. Contemplando retrospectivamente todas las resmas de papel que se han escrito sobre este asunto uno siente que la llegada de la teoría de Einstein sorprendió a personas muy agudas en muy malas condiciones para analizar adecuadamente una cuestión de este género. Sería por cierto contradictorio que, al comparar dos relojes, el tiempo constatado en uno fuese a la vez más y menos que el constatado en el otro. Pero en el caso considerado las comparaciones son (i) entre una serie de, digamos, n lecturas sucesivas de un mismo reloj C' que reposa en $\mathcal{R}'$ con lecturas simultáneas de cada uno de los relojes $C_1,\ldots, C_n$ que reposan en $\mathcal{R}$, por una parte, y (ii) entre una serie de n lecturas sucesivas de un mismo reloj C que reposa en $\mathcal{R}$ con lecturas simultáneas de cada

uno de los relojes C_1',..., C_n' que reposan en $\mathcal{R}'$, por otra; y obviamente no hay ningún motivo para pensar que estas dos series de comparaciones, en el curso de las cuales jamás se compara más de una vez el mismo par de relojes, no puedan arrojar los resultados previstos por la relatividad especial. Las comparaciones se reducen a una sola en el famoso cuento de los mellizos, un niño que pasa toda su vida en el marco inercial de referencia en que nació, y una niña que parte después de nacer en una nave espacial que la transporta a velocidad uniforme cercana a la de la luz en viaje de ida y regreso a una estrella cercana. Según la relatividad especial, a su regreso la mujer será más joven que su hermano varón (suponiendo que los procesos de envejecimiento marchen más o menos al mismo paso que los relojes atómicos).[15] No faltó quien sostuviera que, si el movimiento es relativo, el mellizo tendría que ser *a la vez* mayor y menor que su melliza. Sin embargo, el contraste entre las vidas de los mellizos es mayor aún que la diferencia entre las dos series de comparaciones arriba descritas. Mientras el varón permanece, por hipótesis, atado a un solo marco inercial $\mathcal{R}$, la mujer cambia por lo menos dos veces el marco inercial en que reposa, primero —digamos— saltando de $\mathcal{R}$ a un marco inercial $\mathcal{R}'$ en el que viaja a la estrella, y saltando luego otra vez, al alcanzarla, a un tercer marco inercial $\mathcal{R}''$ en el que retorna a casa. (Doy un ejemplo numérico en el suplemento §8.2, pp. 194–198). Vemos así que las consecuencias controvertidas de la relatividad especial concernientes al tiempo en verdad no contradicen nuestra intuición de las relaciones temporales, sino más bien la manera irreflexiva e inapropiada como solíamos concebirlas y que la relatividad especial nos ha forzado a pensar con más cuidado.

Debo mencionar otra consecuencia desconcertante de la transformación (4) que necesitaremos en la § 2.2. La transformación cinemática de Lorentz no está definida si $v = c$. Si $v > c$, el denominador $\sqrt{1-(v/c)^2}$ es un número imaginario, de modo que la transformación no suministra coordenadas reales. Esta inconsistencia aparente no tiene importancia si los marcos inerciales están asociados por definición a cuerpos masivos ordinarios, pues en la nueva dinámica fundada en los dos principios de la relatividad especial ningún cuerpo

de esta clase puede alcanzar una velocidad $v \geq c$ respecto de otro. La relatividad especial, claro, no excluye que exista otro tipo de materia que se mueve, relativamente a los marcos inerciales, a una velocidad superior a c. Sin embargo, aunque se invirtió algún esfuerzo en descubrir partículas de este tipo, no se las ha detectado jamás. En aras de la simplicidad, doy por sentado que tales partículas no existen.

2.2 Tiempo cíclico e inicio del tiempo en la relatividad general

La relatividad especial se diseñó para calzar con la electrodinámica clásica. Además concuerda muy bien con los resultados experimentales de la física de gran energía. Pero está irremediablemente reñida con la teoría newtoniana de la gravedad, que hacia 1905 era la mejor confirmada de todas las teorías científicas. Muy poco tiempo después de publicar sus primeros trabajos sobre relatividad especial, Einstein había abandonado toda esperanza de acomodar dentro de ella una teoría satisfactoria de la gravedad y a fines de 1907 empezó a trabajar decididamente en una que la desborda. Tras varios intentos fracasados, descubrió ocho años más tarde las ecuaciones del campo gravitacional características de la relatividad general o *ecuaciones de campo de Einstein* (ECE). En los capítulos 1 y 3 comento el curso seguido por su pensamiento hasta encontrarlas. Las siguientes observaciones, numeradas del (1) al (3), serán útiles para quien no haya leído esos capítulos o los recuerde mal.

(1) Einstein se resolvió muy pronto a mantener la identidad entre masa inercial y masa gravitacional introducida por Newton. Infirió de ella que ningún experimento local permite distinguir entre el reposo en un marco inercial de referencia y el reposo en un marco de referencia en caída libre en un campo gravitacional uniforme.[16] Durante toda su búsqueda de una teoría de la grave-

dad Einstein sostuvo este *principio de equivalencia*. Como en el mundo real no hay campos gravitacionales realmente uniformes, el principio de equivalencia implica que será prácticamente imposible hallar un marco de referencia desde el cual definir una carta de Lorentz global (con tiempo de Einstein).

(2) En conferencias públicas dictadas en 1907 y 1908, el matemático Minkowski (1915, 1909) presentó su interpretación geométrica de la cinemática de la relatividad especial. Se la puede explicar sucintamente así. Una carta de Lorentz *x* es una aplicación que asigna un cuádruplo de numeros reales a cada evento.[17] Esto implica que el dominio de *x* —esto es, la colección de todas esas localizaciones potenciales— es una ☆variedad diferenciable☆ cuatridimensional, estructurada por el ☆atlas máximo☆ compatible con la carta de Lorentz *x*. A esta variedad Minkowski la llamó “el mundo” (*die Welt*), pero hoy la llamamos *espaciotiempo*. La nomenclatura de Minkowski sobrevive aún en el término ‘cosmolínea’ (*Weltlinie* en alemán, *worldline* en inglés) con el que nos referimos a una curva☆ en el espaciotiempo cuyo camino☆ comprende los instantes sucesivos de la historia de una partícula material o de un cuanto de radiación o fotón. Minkowski mostró que la acción del grupo de Lorentz (esto es, el grupo generado por las rotaciones espaciales y las transformaciones cinemáticas de Lorentz) determina sobre el espaciotiempo una ☆métrica riemanniana☆ (impropia) $\boldsymbol{\eta}$. Si u^0 es la coordenada temporal y u^1, u^2 y u^3 son las coordenadas espaciales de una carta de U que utiliza el segundo como unidad de tiempo y el segundo-luz (= 299.792.458 metros) como unidad de longitud, los componentes☆ $\eta_{hk} = \boldsymbol{\eta}\left(\frac{\partial}{\partial u^h}, \frac{\partial}{\partial u^k}\right)$ de la métrica $\boldsymbol{\eta}$ relativos a U están dados por $\eta_{00} = 1$, $\eta_{kk} = -1$ si $k > 0$, y $\eta_{hk} = 0$ si $h \neq k$ $(0 \leq h,k \leq 3)$. Una curva espaciotemporal es una curva *temporaloide* si en cada punto p de su camino el vector v tangente a la curva satisface la condición $\boldsymbol{\eta}_p(v,v) > 0$; y es una curva *nula* o *luminoide* si, en todos esos puntos, $\boldsymbol{\eta}_p(v,v) = 0$. Las cosmolíneas de los fotones son

luminoides, mientras que las de las partículas masivas son temporaloides.

(3) Poco después de trasladarse de Praga a Zurich en agosto de 1912, Einstein dio con la idea de que la inexistencia de campos gravitacionales uniformes puede reconciliarse con la validez local de la relatividad especial de un modo análogo al utilizado para reconciliar la curvatura de la tierra con la validez de la geometría euclídea plana a la escala urbana: es cosa de dotar al espaciotiempo con una métrica riemanniana **g** que concuerde con la métrica minkowskiana **η** hasta las cantidades de primer orden en un entorno☆ de cada punto espaciotemporal. Esta nueva geometría jugará un papel protagónico en la teoría de la gravedad si la métrica **g** constriñe a las partículas de prueba en caída libre a seguir determinadas cosmolíneas (geodésicas☆), mientras que ella, a su vez, depende de la distribución espaciotemporal de las fuentes de la gravitación.

Este programa halló su realización en las ECE (reproducidas en la p. 91). Enuncian la igualdad entre un ☆campo tensorial☆ construido con la métrica **g** y sus derivadas primeras y segundas y otro campo tensorial que supuestamente representa todas las fuentes de gravedad (excepto la energía del campo gravitacional mismo). Sin embargo, el segundo campo tensorial es en buena medida una broma,[18] pues somos enteramente incapaces de representar exactamente la distribución de la materia y la energía en el universo y, en el caso de que pudiésemos hacerlo, las ECE presentarían a su lado derecho un campo tensorial de tal complejidad que seríamos incapaces de resolverlas. En la práctica, las ECE se resuelven exactamente solo para campos idealmente simples. Por ejemplo, la solución de Schwarzschild utilizada en la mecánica celeste se refiere al campo esféricamente simétrico de un cuerpo central y las soluciones de Friedmann utilizadas en la cosmología presuponen una distribución de la materia perfectamente isotrópica en torno a cada punto del espacio. Es oportuno destacar que en un espaciotiempo de Friedmann se puede definir una coordenada temporal universal, única salvo por la elección del origen y de la unidad de tiempo: basta considerar simultáneos a todos los

eventos que ocurren sobre la misma hipersuperficie ortogonal a las geodésicas temporaloides. Un resultado similar se obtiene en el espaciotiempo de Schwarzschild utilizando la ortogonalidad con el eje de simetría. Pero, como veremos a continuación, esta notable resurrección del tiempo absoluto se ve más que compensada por la destrucción total del orden universal del tiempo en otros universos ajustados a las ECE.

En 1949, el lógico Kurt Gödel, que era colega de Einstein en el Instituto de Estudios Avanzados de Princeton y solía discutir con él acerca de la realidad o idealidad del tiempo, descubrió una familia de soluciones exactas de las ecuaciones de campo con la propiedad siguiente: un espaciotiempo cuya métrica esté dada por cualquiera de ellas contiene curvas temporaloides cuyos caminos retornan al respectivo punto de partida (Gödel 1949, 1949*b*, 1952). Además, no hay modo de eliminarlas “alzando” el espaciotiempo a un espacio de recubrimiento. Como bien lo expresa Earman: “las curvas no pueden desenrollarse como curvas abiertas en que los eventos se repiten una y otra vez hasta el infinito”, al menos no “sin estropear las propiedades topológicas locales del espaciotiempo” (1995, p. 163). El camino de una de estas curvas temporaloides cerradas es una subvariedad unidimensional homeomórfica☆ al círculo euclidiano y por lo tanto es topológicamente compacto☆. Esta propiedad de las soluciones de Gödel resulta fatal para la representación corriente del tiempo. Localmente —es decir, en un entorno de cada punto— una curva temporaloide cerrada no difiere de la cosmolínea de una partícula material ordinaria, los eventos de cuya historia se ordenan linealmente a lo largo de ella tal como se suceden unos a otros. Pero una curva local cerrada no admite un orden global de sucesión, ya que cada uno de sus puntos a la vez precede y sucede a todos los demás. Además, como es obvio, en un espaciotiempo gödeliano no es posible definir un tiempo universal, ni siquiera por convención.[19] Gödel, claro está, no pretendía hallar una aplicación física para sus soluciones.[20] Advirtió solamente que en un mundo al que se apliquen no puede existir “un trascurso objetivo del tiempo” (1949*a*, p. 558), y arguyó que “la mera compatibilidad con las leyes de la naturaleza de mundos en que no hay un tiempo absoluto discernible y donde, por tanto, no puede existir un

trascurso objetivo del tiempo" no era indiferente al significado del tiempo en nuestro mundo, pues, a menos que uno niegue tajantemente la realidad del tiempo, la cuestión de "si existe o no un trascurso objetivo del tiempo (esto es, de si existe o no un tiempo en el sentido ordinario de esta palabra)" resulta depender "del modo particular como la materia y su movimiento estén dispuestos en el mundo" (1949*a*, p. 562). Por una razón que me escapa, Gödel estimaba que una posición filosófica que tuviese esta consecuencia no podía considerarse satisfactoria.

Posteriormente se han encontrado curvas temporaloides cerradas en otras soluciones de las ECE; pero no hay el menor indicio de que las haya en el mundo real. La discusión al respecto es pues puramente académica, aunque no enteramente ociosa, en cuanto puede ayudarnos a aclarar nuestras ideas sobre asuntos de más enjundia. Hawking y Ellis (1973, p. 189) han señalado que en un universo que contenga curvas temporaloides cerradas sería imposible la ciencia experimental, pues esta requiere que los científicos tengan libertad para ejecutar cualquier experimento, pero una vida entrampada en una curva temporaloide cerrada no goza de ninguna libertad. Sin embargo, no me parece que para un experimentalista aposentado en un universo así sea más difícil eludir el lazo de las curvas temporaloides cerradas, que para nosotros sustraernos a las vastas regiones de nuestro mundo donde las condiciones ambientales son incompatibles con la vida de laboratorio. Antes bien, le será infinitamente más fácil, pues, como la misma observación de Hawking y Ellis nos hace ver, nadie puede *optar* por montarse en una curva temporaloide cerrada: o bien la historia de un objeto engloba esa curva, y entonces ese objeto no puede ser un experimentalista, ni siquiera una persona; o bien definitivamente la deja fuera.

Una objeción menos interesante es que las curvas temporaloides cerradas chocan frontalmente con nuestras intuiciones acerca del tiempo. No veo ninguna razón para que concordasen con ellas. Las curvas temporaloides cerradas son por cierto *temporaloides*, pero este epíteto solo indica una propiedad de sus vectores tangentes. Por cierto, las curvas con vectores tangentes de este tipo se llaman 'temporaloides' porque en la relatividad general sirven para representar las cosmolíneas

de la materia que proveen la base física para el tiempo de nuestras vidas. Pero esto no quiere decir que cualquier curva temporaloide en un espaciotiempo represente la cosmolínea de una partícula material ni mucho menos la vida de una persona. Por último, no hay que perder de vista que las curvas temporaloides cerradas están presentes en ciertos modelos —por lo que sabemos, irreales— de una especie de estructura matemática que tiene *otros modelos* apropiados para representar satisfactoriamente los fenómenos gravitacionales (dentro de un margen admisible pero ciertamente finito de imprecisión).[21] En vista de ello, no hay mucho peligro de que nos quedemos enredados en un ciclo temporaloide de eterno retorno.[22]

En modelos más realistas de la relatividad general surgen al parecer curvas temporaloides compactas de otro tipo. En las exposiciones de la cosmología actual, especialmente en aquellas escritas para el gran público, se habla comúnmente de un primer instante del tiempo, en el cual ocurrió el *Big Bang* y empezó la evolución del universo. La cosmología del *Big Bang*, respaldada observacionalmente por la recesión de las galaxias descubierta por Hubble (1929) y por la casi perfectamente isotrópica radiación térmica de fondo descubierta por Penzias y Wilson, tiene su cimiento teórico en la relatividad general, específicamente en las soluciones de Friedmann de las ECE, arriba aludidas.[23] Los modelos de Friedmann en expansión representan universos con *Big Bang*, idealizados. En tales modelos, la densidad de la materia ρ es constante en cada hipersuperficie ortogonal a las geodésicas temporaloides y decrece sin cesar a lo largo de estas geodésicas. Si el tiempo se mide con relojes estándar a lo largo de las geodésicas temporaloides, una gráfica de ρ contra el tiempo mostrará que ρ disminuye en un plazo finito desde un valor que supera cualquier cantidad asignable hasta su presente valor finito. Dicho con más precisión: si la coordenada temporal definida por relojes estándar sobre cada geodésica temporaloide aumenta en la dirección del futuro y le asignamos la coordenada 0 al momento actual, la densidad ρ aumenta más allá de todo límite cuando la coordenada temporal se aproxima, desde arriba, a cierto número finito negativo $-T$.

Las geodésicas en cuestión son las cosmolíneas de la materia en el modelo de Friedmann. Los libros de divulgación científica sugie-

ren francamente que cada una de ellas debe considerarse como una aplicación del intervalo cerrado $[-T,0] \in \mathbb{R}$ sobre una subvariedad unidimensional compacta del espaciotiempo. Tales aplicaciones pueden extenderse hacia el futuro, a aplicaciones definidas sobre algún intervalo cerrado-abierto $[-T,\varepsilon)$, donde ε es un número real positivo pequeño, pero no hacia el pasado, a aplicaciones sobre el intervalo abierto $(-(T+\varepsilon),\varepsilon)$.

Pero esto es un disparate. Un espaciotiempo relativista no contiene puntos donde la densidad de la materia sea infinita y por ende la métrica no esté definida. Por lo tanto, aunque según las condiciones que hemos supuesto la coordenada temporal se aproxima desde arriba a $-T$ indefinidamente a lo largo de cada cosmolínea de la materia, no hay ningún punto sobre ninguna de estas cosmolíneas al que corresponda la coordenada $-T$. En efecto, si E es un evento cualquiera, todo lo temprano que se quiera, en un universo de Friedmann en expansión, la densidad ρ es finita en E y hay en ese universo otros eventos aún más tempranos que E en que ρ es mayor. El número T es la cota inferior máxima de la coordenada temporal establecida del modo indicado, pero no forma parte del recorrido de esa coordenada.

Los matemáticos conocen métodos para agregarle puntos fronterizos a ☆variedades riemannianas☆. Apelando a alguno de ellos, el cabo suelto en el pasado de cada geodésica temporaloide de un modelo expansivo de Friedmann puede cerrarse por decreto. Pero esta clausura no es exigida por la teoría ni está respaldada por experimentos u observaciones. Según Kant la cosmología está afligida necesariamente por la siguiente antinomia: o bien el universo tuvo un comienzo en el tiempo y tenemos que preguntarnos qué había antes, o bien el universo no tuvo comienzo y tenemos que admitir la idea, que él consideraba absurda, de un pasado infinito que sin embargo ahora mismo está completo. Gracias a la geometría diferencial moderna, la relatividad general escapa a ambos cuernos del dilema. El universo del *Big Bang* no tiene comienzo, pero hasta la fecha ha durado solo por un tiempo finito.[24]

3

Gravedad y curvatura

La teoría general de la relatividad de Einstein o *relatividad general* concibe los fenómenos gravitacionales como manifestaciones de la curvatura del espaciotiempo en que ocurren los eventos físicos. En este capítulo bosquejo el curso de pensamiento que llevó a Einstein a concebirlos así y examino brevemente las propuestas de Harold Jeffreys y Richard Feynman para retener las ecuaciones del campo gravitacional de Einstein desechando su interpretación geométrica.

3.1 Breve historia de la gravedad

Desde la publicación de las ecuaciones de campo de Einstein (ECE) en noviembre de 1915 (Einstein 1915*i*), suele hacerse la pregunta "¿Por qué hay que concebir la gravedad como curvatura del espaciotiempo?" Se la puede entender así: existe en el mundo, independientemente de nuestros modos de pensar, una fuerza natural bien definida que llamamos 'gravedad', o una clase natural de efectos que llamamos 'gravitacionales', tales que, cuando nos ponemos a pensar acerca de ellos, tenemos que concebirlos como consecuencia del hecho de que el mundo en que vivimos es una realización de la estructura matemática que llamamos 'espaciotiempo curvo'; la pregunta inquiere por qué hay que concebirlos así y, más específicamente, qué resultados de la investigación científica nos constriñen a pensar de esta manera.

Esta interpretación, aunque aparentemente natural, implica sin embargo, que la pregunta es trivial o impertinente. El término 'gravedad' es la forma castellana de la voz latina *gravitas*, esto es, 'pesadez'. Su extensión ha variado drásticamente más de una vez con sucesivos cambios en la física. La familia de los fenómenos gravitacionales comprendía en la antigüedad solo la caída de los cuerpos

pesados sobre la tierra y la presión que ejercen hacia abajo cuando se les impide caer; pero Newton la amplió incluyendo en ella el movimiento de los astros. La teoría newtoniana de la atracción universal todavía determina el significado de 'gravedad' y 'gravitación' en el castellano corriente, pero cuando concebimos la gravedad como curvatura del espaciotiempo gobernado por las ECE, la familia de los fenómenos gravitacionales se amplía *ipso facto* y pasa a incluir también otros a primera vista tan heterogéneos como la variación en el movimiento de púlsares binarios debido a la emisión de ondas gravitacionales, la aparente duplicación de fuentes luminosas debido a la presencia de lentes gravitacionales y el lento enfriamiento de la radiación térmica de trasfondo descubierta por Penzias y Wilson. La ambigüedad del término 'gravedad' escinde en muchas preguntas lo que parecía ser una sola: "¿Por qué hay que concebir la gravedad como curvatura del espaciotiempo?" Ni la gravedad aristotélica ni la newtoniana se conciben como curvatura del espaciotiempo, de modo que, aplicada a ellas, la pregunta es impertinente. Por otra parte, la teoría de Einstein entiende los fenómenos gravitacionales precisamente como manifestaciones de la curvatura del espaciotiempo, y por tanto no puede llamarnos la atención que la gravedad einsteiniana tenga que concebirse así. En este caso, pues, la pregunta es trivial.

La pregunta admite, sin embargo, otra interpretación que no suscita esta doble crítica. Podemos parafrasearla como sigue: ¿por qué los fenómenos físicos que la teoría de Einstein clasifica como gravitacionales deben considerarse tales en la acepción en que Einstein usa este término? ¿por qué hemos de juzgarlos como pertenecientes a una clase natural y juzgarlos así precisamente por la razón que da la relatividad general, a saber, que cada uno de ellos manifiesta a su modo la presencia de curvatura espaciotemporal? Adviértase que cuando Einstein abordó el estudio de la gravedad entendía referirse a los fenómenos reunidos bajo el concepto de Newton. Otros fenómenos, como el llamado retardo gravitacional de los relojes y la deflección de la luz de una estrella que pasa bordeando al sol, y también los que mencioné antes, se le fueron agregando a la extensión del concepto de Einstein en virtud de su teoría que predecía tales fenómenos e indujo a los investigadores a buscar algunos de ellos. Vista bajo esta

perspectiva, nuestra pregunta tiene que ver con el cambio conceptual: ¿por qué juzgó Einstein necesario reconcebir los fenómenos comprendidos bajo el concepto newtoniano de gravedad y entenderlos como efectos de la curvatura del espaciotiempo? Pero también es una cuestión de idoneidad: ¿por qué el concepto einsteiniano de gravedad y la clasificación que impone a los fenómenos deben preferirse a las otras alternativas disponibles? ¿qué nos obliga o siquiera nos invita a pensar que este concepto se ajusta a los fenómenos que intenta captar? Aquí nos ocuparemos solo con la primera cuestión y dejaremos enteramente de lado la otra. (El suplemento §8.3, en las pp. 198–201 reproduce un texto inconcluso donde abordo superficialmente la segunda cuestión).[1]

3.2 Geometría y caída libre

Einstein empezó a trabajar en una teoría de la gravedad en 1907. ¿Qué lo indujo a ello? En mi opinión, la respuesta a esta pregunta es muy simple: La ley de gravedad de Newton está en conflicto frontal con los principios generales de la física postulados por Einstein en 1905, los principios en que descansa la relatividad especial. Como se vio en la §1.2 (pp. 17 y ss.), Einstein (1905*r*) postuló la completa equivalencia de todos los marcos inerciales de referencia y la constancia de la velocidad de la luz en el vacío con respecto a cualquiera de esos marcos y demostró que la afirmación conjunta de estos dos postulados implica la invariancia de las leyes de la física bajo las transformaciones del grupo de Lorentz. La ley de gravedad de Newton no es invariante bajo dichas transformaciones y por tanto tenía que ser revisada para que tuviera cabida en la nueva física. Poincaré (1906) lo señala en un artículo en que llegó por su propia cuenta a conclusiones prácticamente indistinguibles de las de Einstein.[2] En la sección final de este trabajo, Poincaré bosqueja una teoría de la gravedad que es invariante bajo las transformaciones de Lorentz. No voy a explicar esta teoría aquí, pero mencionaré dos pasos previos que dio Poincaré antes de formularla.

En primer lugar, para abordar el problema de la gravedad, representó los fenómenos físicos en una ☆variedad cuatridimensional☆ representada mediante coordenadas $\langle x,y,z,t \rangle$ (coordenada temporal t expresada en segundos, coordenadas espaciales x, y, z expresadas en segundos-luz) y probó que la siguiente forma cuadrática

$$x^2 + y^2 + z^2 - t^2 \tag{1}$$

es invariante bajo las transformaciones de Lorentz. Como expliqué en la §1.3 (pp. 23 y ss.), este punto de vista fue pronto adoptado por Minkowski, cuya lectura cronogeométrica de la relatividad especial fue decisiva para la formulación de la teoría de la gravedad de Einstein o relatividad general.

En segundo lugar, Poincaré resolvió que, entre las distintas formas posibles de una ley de gravedad invariante bajo las transformaciones de Lorentz, debía elegir aquella que —en situaciones en que los cuerpos en interacción gravitacional se mueven unos respecto a otros lentamente— difiriese menos de la ley de Newton. Porque, dice, “no parece que las observaciones astronómicas indiquen ninguna diferencia significativa con la ley de Newton” (1906, p. 167). Este comentario demuestra que Poincaré, la máxima autoridad mundial en materia de mecánica celeste, no consideraba significativa la discrepancia entre el valor previsto y el valor observado del avance secular del perihelio de Mercurio (igual a casi ocho partes en mil).

La actitud de Poincaré contrasta con la de Einstein, quien comunicó el 24 de diciembre de 1907 a Conrad Habicht que estaba trabajando en una nueva teoría de la gravedad de la que esperaba que diese cuenta de “las variaciones todavía inexplicadas” del perihelio de Mercurio (CP, vol. 5, p. 82). Este mensaje navideño de Einstein respalda al parecer la familiar tesis empirista, según la cual cada avance de la física teórica está motivado lisa y llanamente por el fracaso de alguna predicción de la teoría vigente. El interés de Einstein en la anomalía del perihelio de Mercurio está confirmado por el hecho de que cuando finalmente pudo resolverla utilizando la teoría de la gravedad que publicó el 11 de noviembre de 1915 (Einstein 1915*g*), se apresuró a publicar su solución una semana más tarde (1915*h*), antes de estar seguro siquiera de que esa teoría merecía desarrollarse.[3] No

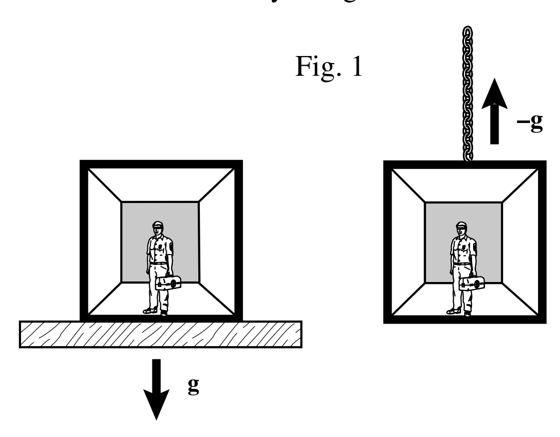
obstante, pienso que para Einstein la anomalía del perihelio de Mercurio era solo un test fácilmente aplicable a la teoría de la gravedad que él buscaba motivado por intereses mucho más amplios. Era una buena cosa, claro está, que la nueva teoría superase a la de Newton en este respecto, aunque fuera un asunto menor;[4] pero Einstein estaba dispuesto a adoptar una que no tuviera esta virtud. Así, no titubeó en publicar en 1913 la teoría de la gravedad que elaboró junto con Marcel Grossmann y que no resuelve la anomalía de Mercurio (Einstein y Grossmann 1913). Más aún, cuando presentó esta teoría a la asamblea de científicos alemanes reunida en Viena ese mismo año, rindió homenaje a Newton en términos aún más vigorosos que Poincaré: la ley newtoniana "de la interacción entre dos puntos-masa que gravitan [...] ha demostrado ser tan exactamente correcta que, desde el punto de vista de la experiencia, no hay ninguna razón decisiva para dudar de su rigurosa validez" (1913*c*, p. 1249). Como quiera que se piense sobre el impulso motivador que Einstein pueda haber recibido de la anomalía de Mercurio, esta no lo ha guiado en su búsqueda. Einstein no ensayó correcciones que calzasen mejor con las observaciones,[5] sino que fue derecho tras un principio universal capaz de ponernos sobre la pista de la raíz misma de la gravedad. El principio que halló y llamó *principio de equivalencia* estaba incluido también, aunque solo como corolario, en la teoría de Newton; pero Einstein supo blandirlo de una manera conducente a una concepción nueva —y enteramente inesperada— de los fenómenos gravitacionales.

Para entender lo que Einstein hizo de este principio hay que recordar los rasgos centrales de la concepción newtoniana. En virtud de la primera ley del movimiento de Newton, un cuerpo reposa o se mueve en línea recta y a velocidad constante a menos que una fuerza externa lo acelere; en virtud de la segunda ley, la aceleración que un cuerpo recibe de una fuerza externa es proporcional a esta e inversamente proporcional a la cantidad de materia o *masa* que el cuerpo contiene. Por otra parte, desde Galileo la física moderna sostiene que la aceleración de gravedad en un lugar dado de la tierra es la misma para todos los cuerpos. Newton aceptó esta proposición e incluso la verificó experimentalmente con péndulos hechos de distintos materiales, que según él le permitieron confirmarla con un margen de error

inferior al 1 por mil (1726, p. 400). Newton resolvió el aparente conflicto con sus leyes del movimiento postulando que el efecto de la fuerza gravitacional sobre cualquier cuerpo es directamente proporcional a su masa. De este modo, la *inercia* con que el cuerpo resiste a la aceleración de gravedad es compensada exactamente por su susceptibilidad a gravitar, esto es, por su *carga pesada*.[6]

En el contexto de la relatividad especial este manejo del asunto se torna sumamente problemático. La inercia de un cuerpo no es una cantidad invariante bajo las transformaciones de Lorentz y varía con su velocidad relativa al marco inercial al que referimos su estado de movimiento. ¿Vale otro tanto para la carga gravitacional? Max Planck sospechaba que no y que la igualdad empíricamente constatada entre masa "inercial" y masa "ponderable" no pasaba de ser una aproximación útil (1907, p. 544). Por su parte, Einstein decide más o menos al mismo tiempo atenerse al resultado nulo de los experimentos de Newton con los péndulos, cuya precisión Eötvös (1889) había mejorado mediante una balanza de torsión, reduciendo el error a 1 parte en 100 millones.[7] Sin embargo, Einstein no lo explica por la autocompensación de una magnitud física que actúa contra sí misma, sino por una extensión del principio de relatividad, que él juzgaba natural. No en vano la versión inicial de este principio, restringida a los marcos inerciales, le permitía dar cuenta del resultado nulo del experimento de Michelson y Morley sin apelar a efectos físicos compensatorios como la contracción de Lorentz y Fitzgerald.[8]

Fig. 1



El principio de equivalencia se enuncia usualmente en los escritos de Einstein como una aplicación de la relatividad a marcos de referencia uniformemente acelerados. Tales enunciados pueden parafrasearse así (véase la Fig. 1):

> *No hay modo de determinar mediante experimentos físicos ejecutados en un laboratorio cerrado, si ese laboratorio reposa en un campo gravitacional uniforme* **g** *o si se mueve con aceleración uniforme –***g** *relativamente a un marco de referencia inercial.*

Si esto es así (véase Fig. 2),

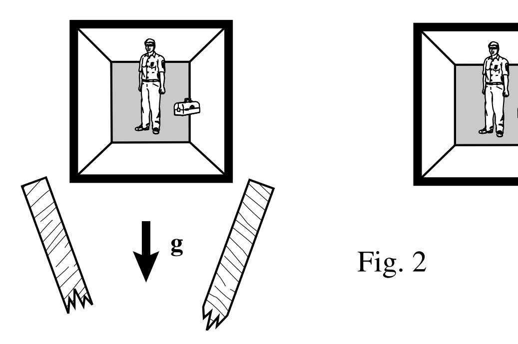


Fig. 2

> *No hay modo de determinar mediante experimentos físicos ejecutados en un laboratorio cerrado, si ese laboratorio cae libremente en un campo gravitacional uniforme o si se mueve inercialmente —esto es, a velocidad constante— en el espacio libre de campos gravitacionales.*

Esta segunda versión del principio de equivalencia de Einstein está implícita en el corolario VI de las leyes del movimiento de Newton, que podemos traducir así:

> Si cuerpos que se mueven unos respecto a otros de cualquier modo son urgidos en direcciones paralelas por fuerzas aceleradoras iguales, seguirán moviéndose unos respecto a otros del mismo modo que si no hubieran sido urgidos por esas fuerzas.
>
> (Newton 1726, p. 21)

También cuando Einstein relata cómo se le ocurrió el principio de equivalencia sugiere la segunda versión más bien que la primera:

> En 1907, cuando trabajaba en un artículo panorámico sobre la teoría especial de la relatividad para el *Jahrbuch der Radioaktivität und Elektronik*, tuve también que tratar de modificar la teoría newtoniana de la gravedad de tal modo que sus leyes concordaran con mí teoría. Los intentos hechos en este sentido mostraban que la empresa era realizable, pero no me satisfacían porque tenían que apoyarse en hipótesis físicamente infundadas. Me vino entonces la idea más feliz de mi vida en la forma siguiente. El campo gravitacional sólo tiene una realidad relativa, tal como el campo eléctrico generado por inducción magneto-eléctrica. *Porque para un observador que cae libremente desde el techo de una casa no hay durante su caída* —al menos en su entorno inmediato— *ningún campo gravitacional*. Si el observador se desprende de algunos objetos, estos permanecen relativamente a él en estado de reposo o de movimiento uniforme, independientemente de la particular naturaleza química o física (esta consideración ignora, por cierto, la resistencia del aire). El observador tiene derecho pues a interpretar su estado como 'reposo'.
>
> (Einstein, CP, vol. 7, p. 265)

Esta fantasía de Einstein de un observador en caída libre concuerda muy bien con la visión ahora familiar de los astronautas en órbita. No importa cuál sea su composición, un cuerpo que cae libremente en un campo gravitacional uniforme no exhibe signo alguno de tensión o apremio, sino que flota en un estado de ingrávido sosiego, tal como si se moviese inercialmente a una distancia infinita de todas las fuentes de gravedad.

Del principio de equivalencia y el principio de conservación de la energía Einstein dedujo que la luz gravita y que los relojes naturales tienen que atrasarse cuando se los traslada a un sitio donde el potencial gravitacional es más alto. Ambos fenómenos han sido luego observados. Misner, Thorne y Wheeler (1973, pp. 187–189) arguyen —siguiendo a Schild, dicen— que si los relojes sufren dicho atraso, el espaciotiempo no puede ser minkowskiano, o sea, plano, de modo que aun la presencia de un campo gravitacional uniforme es un signo de curvatura. No creo que su argumento sea correcto, pero es instruc-

tivo. Lo esbozaré y luego presentaré mi crítica. Considérese una fuente de luz en reposo en un campo gravitacional uniforme. Cada segundo emite una señal luminosa en la dirección en que aumenta el potencial. La recepción de las señales se registra en un punto de potencial más alto, fijo también en el campo. Debido al retraso gravitacional de los relojes, las señales llegan a intervalos mayores que un segundo. Sean *A* y *B*, respectivamente, la emisión y la recepción de un señal, y *C* y *D* la recepción y emisión de la próxima (Fig. 3). Los eventos *A*, *B, C* y *D* ocupan entonces los vértices de un paralelógramo en el espaciotiempo, con lados opuestos desiguales *AC* y *BD.* En un espaciotiempo plano no puede haber un paralelógramo con esta propiedad. Por lo tanto, el espaciotiempo que subyace a un campo gravitacional uniforme es curvo. El argumento valdría también si el campo gravitacional deflecta la luz, con tal que las señales sucesivas describan trayectorias espaciotemporales curvadas del mismo modo (Fig. 3, derecha); pues en tal caso los eventos *A*, *B, C* y *D* siguen ocupando los vértices de un paralelógramo con lados opuestos desiguales.

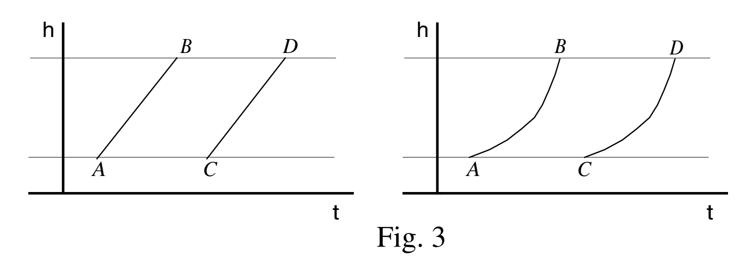


Fig. 3

Pero —y aquí viene mi crítica— el argumento supone que la geometría del espaciotiempo se refleja en las mediciones efectuadas con varas rígidas y relojes, no importa cuál sea el estado de movimiento de estos instrumentos. Sin embargo, si la geometría es minkowskiana, tales mediciones la reflejarán si y solo si las varas y relojes reposan en un marco de referencia inercial. Como ese no es el caso bajo las condiciones prescritas para nuestro experimento mental, el resultado no implica que el espaciotiempo subyacente sea curvo y no minkowskiano. Lo que sí demuestra, en cambio, es que los sistemas de coordenadas espaciotemporales definidos del modo habitual

sobre un marco de referencia que reposa en la superficie de la tierra no pueden tener el significado métrico que tradicionalmente se les atribuye. Einstein nos cuenta que el progreso de su investigación sobre la gravedad se vio dificultado por su resistencia a aceptar esta conclusión. Pues, como observa en su autobiografía, "uno no se libra tan fácilmente de la idea de que hay que atribuir a las coordenadas un significado métrico inmediato" (Einstein, 1949, p. 66).[9]

Con todo, a la larga desechó la exigencia tradicional de que las coordenadas físicas contengan información geométrica y cronométrica. Einstein (1916*e*, pp. 774–775) asocia expresamente este paso esencial a su análisis de las mediciones practicadas en un disco giratorio mediante varas y relojes fijos en él (véase el suplemento §8.1, pp. 191–194). Había aludido a esto, de paso, ya en 1912. En ese mismo año dio otro paso, que estimo mucho más importante, y que en sus escritos suele aparecer vinculado —algunos dirán confundido— con aquel. Me refiero a la decisión de representar el campo gravitacional en un espaciotiempo que se supone localmente —en rigor, *tangencialmente*— minkowskiano, pero cuya métrica global no se postula a priori sino que ha de determinarse observando los fenómenos de la gravedad. Esta decisión no puede basarse en el solo principio de equivalencia, pero seguramente se apoya en él tomado conjuntamente con el comportamiento efectivo de los campos gravitacionales existentes.

Consideremos un sistema de coordenadas con significado métrico del tipo que en la p. 51 acordamos llamar *cartas de Lorentz* y que normalmente se utiliza en la relatividad especial. Tal sistema consta de coordenadas cartesianas definidas mediante varas rígidas en reposo en un marco inercial de referencia y una coordenada temporal definida mediante relojes repartidos por todo el marco inercial y sincronizados según el método de Einstein (1905*r*, §1).[10] Adoptamos unidades de longitud y de tiempo tales que la velocidad de la luz en el vacío sea igual a 1.[11] El espaciotiempo minkowskiano de la relatividad especial admite cartas de Lorentz globales justamente porque su curvatura es igual a 0 por doquier. A la luz del principio de equivalencia en su segunda versión es claro que también es posible definir una carta de Lorentz global sobre un campo gravitacional

uniforme que cubra todo el universo. De esto se infiere inmediatamente que no puede ser válido el argumento atribuido a Schild que Misner, Thorne y Wheeler invocan a favor de la curvatura del espaciotiempo en presencia de gravitación uniforme. Pero también se infiere que el espaciotiempo no puede ser plano en el mundo real. Pues los campos gravitacionales existentes son solo aproximada y localmente uniformes. Como es bien sabido, al movernos sobre un meridiano la dirección de la aceleración de gravedad varía con la latitud hasta completar una rotación completa de 360°. Por lo tanto, el espaciotiempo en que vivimos no admite una carta global de Lorentz, de modo que no es minkowskiano, sino curvo.

Este razonamiento sugiere que hay un vínculo entre la métrica del espaciotiempo y la gravedad, pero siendo puramente formal y negativo, no arroja ninguna luz sobre ese vínculo. Creo que fue al encarar esta sugerencia que entró en juego la tremenda inventiva de Einstein o, si se quiere, su extraordinaria capacidad para ver analogías y afinidades que no se le ocurrían a otros. En la relatividad especial hay una conexión obvia entre la geometría del espaciotiempo y el movimiento inercial. En ausencia de fuerzas externas, la cosmolínea (p. 32) de una partícula masiva ordinaria no cambia de dirección; por ello, las coordenadas asignadas a tal partícula libre en una carta de Lorentz satisfacen una ecuación lineal y la representación gráfica apropiada de su cosmolínea es una línea recta. Esto se deja expresar con más precisión así: en cualquier instante de su historia, una partícula masiva tiene una cosmovelocidad bien definida, representada por la tangente☆ a su cosmolínea en ese instante; si la partícula no es urgida por fuerzas externas, todas las tangentes a su cosmolínea son, a lo largo de esta, paralelas entre sí.[12] Se podría decir que la partícula, a falta de campos de fuerza que la impulsen en una u otra dirección, se deja guiar por la geometría del espaciotiempo. Con singular audacia, Einstein concibe la gravedad según esta analogía. Había establecido la equivalencia entre el movimiento inercial y la caída libre en un campo gravitacional uniforme. Hacía mucho que simpatizaba con la sugerencia un tanto nebulosa de Mach, según la cual la inercia, no menos que la gravedad, se debe a la distribución de la materia en el universo. Ahora vio la oportunidad de darle a esta sugerencia una

forma inteligible y operante. Se decidió a ver la geometría del espaciotiempo como la única guía de las partículas en caída libre en un campo gravitacional cualquiera y a subordinar la configuración efectiva de esa geometría a la distribución universal de la materia.

Einstein regresó de Praga a Zürich en agosto de 1912. Años más tarde recordaba haberse percatado en ese momento de la analogía entre el problema matemático que afrontaba la teoría de la gravedad que tenía en mira y la teoría de las superficies curvas de Gauss.[13] Gauss (1828) introdujo un modo de describir las relaciones geométricas sobre una superficie *intrínsecamente*, esto es, en términos de los puntos y líneas que hay sobre ella, sin referirse al espacio en que está incrustada. El mismo incluye nociones intrínsecas de distancia y curvatura y ecuaciones diferenciales para las curvas más rectas —las *geodésicas*☆— de la superficie. La geometría intrínseca de una superficie con curvatura gaussiana distinta de 0 es, por cierto, no euclídea; pero cada punto de la superficie tiene un entorno en que su geometría intrínseca concuerda bien —hasta las cantidades de primer orden— con la geometría del plano euclídeo. Ello hace posible, por ejemplo, que el centro de muchas ciudades hispanoamericanas parezca dividido en manzanas cuadradas, aunque es imposible trazar un cuadrado genuino sobre la superficie curva de la tierra. De un modo análogo, Einstein podía dar cuenta del éxito local de la relatividad especial en nuestros laboratorios, postulando que en un entorno apropiado de cada evento puntual la geometría del espaciotiempo concuerda hasta las cantidades de primer orden con la geometría plana de Minkowski. Esto implica que el espaciotiempo es una ☆variedad riemanniana☆ con la misma signatura☆ que el espaciotiempo de Minkowski.[14] Por lo mismo, la clasificación minkowskiana de las curvas en temporaloides☆, espacialoides☆ y nulas☆ se puede extender sin dificultad a un espaciotiempo curvo. En el nuevo contexto, la analogía entre movimiento inercial y caída libre opera así: se sobrentiende que una partícula en caída libre describe una geodésica temporaloide, de modo que su cosmovelocidad permanece paralela a sí misma a lo largo de su cosmolínea o, en otras palabras, que las partículas en caída libre no experimentan cosmoaceleración alguna. En las variedades riemannianas, la presteza con que una geodésica temporaloide converge

hacia las curvas vecinas de la misma clase o diverge de ellas depende directamente de la curvatura local.[15] Así, la mutua atracción aparente entre cuerpos vecinos bajo la influencia de la gravedad resulta ser, desde este punto de vista, una manifestación de la curvatura del espaciotiempo. De este modo se aclara finalmente el enigma del fundamento o razón de ser de la gravitación universal y entendemos por qué el mismo Newton no fue capaz resolverlo.[16] Newton trató de hallar la causa que desvía los cuerpos de su movimiento natural uniforme en línea recta y los fuerza a caer unos sobre otros, pero ocurre simplemente que no hay tal cosa: la caída libre no es más forzada que el movimiento inercial y su curso no resulta de la intervención de un agente.

En agosto de 1912 Einstein ya había oído hablar de la generalización de los métodos de Gauss a espacios de cualquier dimensión finita propuesta por Riemann (1854). Su amigo Marcel Grossmann le señaló la herramienta matemática apropiada para representar los fenómenos físicos en una variedad riemanniana cualquiera: el cálculo diferencial absoluto de Ricci y Levi-Civita (1901). Blandiendo este poderoso instrumento, Einstein y Grossmann (1913) produjeron "una teoría generalizada de la relatividad y una teoría de la gravedad" que comparte con la relatividad general tanto su propósito como muchos rasgos esenciales. Las principales ideas comunes a ambas teorías pueden resumirse así:

1. El espaciotiempo es una variedad riemanniana cuatridimensional cuya métrica —☆un campo tensorial simétrico y no degenerado de rango 2☆— concuerda hasta las cantidades de primer orden con la métrica de Minkowski en un entorno de cada evento. Los relojes naturales miden el tiempo propio a lo largo de su respectiva cosmolínea.

2. El campo métrico guía la caída libre tal como la métrica de Minkowski guía el movimiento inercial en la relatividad especial. Una partícula masiva ordinaria sin momento angular y libre de toda influencia externa salvo la gravedad describe una geodésica temporaloide. Los fotones trazan geodésicas nulas.

3. El campo métrico es, pues, la misma cosa que el campo gravitacional y depende por tanto de la distribución de la materia y la energía no gravitacional en el espaciotiempo.
4. La dependencia entre el campo métrico y la materia-energía se rige por las ecuaciones del campo gravitacional, un sistema de ecuaciones diferenciales de segundo orden cuyas incógnitas son los componentes de la métrica, que concuerdan con la ecuación de Poisson de la teoría newtoniana en el caso límite de cuerpos que se mueven lentamente en un campo gravitacional débil.

De este modo y por las razones expuestas, Einstein había concebido ya en 1913 la gravedad como curvatura. Faltaba aún un sistema satisfactorio de ecuaciones diferenciales que vinculara la curvatura del espaciotiempo a la distribución de la materia. Las ecuaciones de campo de la relatividad general —las ECE— se publicaron el 25 de noviembre de 1915 (Einstein 1915*i*). Como hay razones muy poderosas para pensar que son la única solución viable del problema matemático descrito bajo el número 3, muchos se preguntan por qué Einstein no las halló de inmediato y se afanó durante más de dos años en defender las ecuaciones viciadas de la teoría de Einstein y Grossmann. Pero este no es un tema que podamos tratar aquí.[17]

3.3 Gravedad sin geometría: Jeffreys y Feynman

Me referiré, en cambio, a una cuestión puramente conceptual que ha estado en el tapete casi tanto tiempo como la propia relatividad general. Es una cuestión que generalmente se ignora en los manuales, pero que he visto asomar en dos publicaciones filosóficas recientes, una de las cuales (Kaiser 1995) habla largamente sobre el libro de Richard P. Feynman (1995) que luego comentaré, mientras que la otra (Mayo 1996, pp. 287–288, 291–292) cita con aprobación un artículo de Harold Jeffreys (1919)[18] que planteó la cuestión cuando las mediciones practicadas por Eddington y sus compañeros durante el

eclipse total de sol de ese año supuestamente confirmaron la teoría de la gravedad de Einstein.

Para simplificar, supongamos que los fenómenos que la física actual llama gravitacionales constituyen una clase natural y que las ECE resumen nuestro conocimiento de esos fenómenos en la mejor forma posible ahora. En estas ecuaciones, las incógnitas son diez funciones escalares que son los componentes de un campo tensorial simétrico covariante de rango 2, relativos a una carta☆ del espaciotiempo.[19] Para un matemático, ese campo tensorial en nada se distingue de una métrica riemanniana que concuerda localmente hasta las cantidades de primer orden con la métrica de Minkowski. Pero cabe siempre preguntar ¿significa esta perfecta equivalencia matemática que *una solución de las ecuaciones de Einstein* obtenida bajo supuestos físicos realistas efectivamente caracteriza *la geometría del universo*? Esta es la cuestión conceptual que quiero examinar. Dudo que pudiera ocurrírsele a quien tuviese presente el proceso descrito en la sección anterior, al cabo del cual Einstein descubrió sus ecuaciones de campo no como un conjunto de marcas de tinta negra sobre una hoja blanca de papel, sino como la expresión simbólica adecuada de cierto modo de pensar (que, como vimos, estaba entero impregnado de geometría). Pero los filósofos no siempre prestan atención a este género de procesos, que no pertenecen al epistémicamente sublime "contexto de la justificación", sino al humilde y contingente "contexto del descubrimiento". Y la investigación científica normalmente los olvida.

Jeffreys (1919) entiende la relatividad general de un modo decididamente ageométrico. Caracteriza, es cierto, las cosmolíneas de las partículas en caída libre como aquellas curvas sobre las cuales alcanza un valor extremo la integral curvilínea cuyo integrando *ds* está definido —como el ☆elemento lineal☆ de una variedad riemanniana— por la relación

$$ds^2 = \sum_i \sum_j g_{ij} dx_i dx_j \qquad (2)$$

donde los g_{ij} son ciertas funciones de las coordenadas. Pero Jeffreys nos previene expresamente que

> No hemos hecho ninguna suposición respecto a la identidad de *ds* con el elemento de línea en el espacio-tiempo cuatridimensional, o que vincule los g_{ij} con la "curvatura" de tal espacio-tiempo. *Espacio, tiempo, y curvatura son nociones metafísicas cuya utilidad, si tienen alguna, depende enteramente de las peculiaridades individuales de la persona que trata de entenderlas.* Las cantidades son aquellas que se derivan de mediciones físicas, de suerte que la invariancia de *ds* ha dejado de ser axiomática.
>
> (Jeffreys 1919, p. 146; cursiva mía)

Así pues, según Jeffreys, el integrando *ds* de la integral extremalizada sobre las cosmolíneas de las partículas en caída libre no tiene nada que ver con la geometría —o la cronogeometría— del universo.[20]

La lectura ageométrica de la teoría de la gravedad de Einstein por Richard Feynman es bastante más ilustrada y por ello mucho más interesante que la de Jeffrey. En sus *Lecciones sobre la gravitación* de 1962/63, Feynman (1995) juega con la idea de que se podría avanzar de una teoría de la gravedad clásica a una cuántica de un modo análogo a como Tomonaga, Schwinger y el propio Feynman procedieron con éxito de la electrodinámica clásica a la cuántica (cf. Schwinger 1958, Schweber 1994). En la primera lección, narra el siguiente episodio de una historia fabulosa de la ciencia: los físicos de Venus conocen bien unos treinta campos físicos diferentes y han desarrollado teorías idóneas respecto de todos ellos, pero solo últimamente han descubierto la gravedad; la conciben naturalmente como un trigésimo primer campo de fuerzas, por analogía con los otros treinta. Feynman considera qué clase de partícula irá asociada a este campo. De las propiedades familiares de los fenómenos gravitacionales, especialmente aquellas expresadas por el principio de equivalencia y la ley del cuadrado inverso de la distancia, concluye que esta partícula no puede ser ninguna de las que ya conocemos —en particular, no puede ser el neutrino— , sino que es una partícula hasta ahora desconocida de spin 2 que Feynman llama *gravitón*. El campo asociado a una partícula de spin 2 tiene que ser descrito por un potencial tensorial simétrico de rango 2 que designaré con **h**. Siguiendo el ejemplo de la electrodinámica cuántica, el campo tensorial **h** se define sobre "el espacio de la relatividad especial" (Feynman *et al.* 1995, p. 112), esto

es, el espaciotiempo de Minkowski. Como sabemos, la estructura métrica del espaciotiempo de Minkowski está dada por otro campo tensorial simétrico de rango 2, que designo como es habitual con $\boldsymbol{\eta}$. Feynman esboza una teoría cuántica del campo gravitacional **h**. Para calcular sus efectos macroscópicos, por ejemplo, las órbitas de los planetas que circulan alrededor de una estrella, esta teoría cuántica tiene que reducirse a su forma clásica. Esto se efectúa derivando la teoría clásica de un principio variacional que envuelve una integral del lagrangiano. Mediante un elegante argumento, Feynman deduce la siguiente fórmula para esta integral:

$$m_0\left[-\frac{1}{2}\int d\tau\, g_{\mu\nu}(x)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\right] \tag{3}$$

donde τ es el tiempo propio y los $g_{\mu\nu}$ son los componentes relativos a la carta x de un campo tensorial **g**, simétrico y de rango 2, introducido en aras de la brevedad por la definición siguiente:

$$\mathbf{g}(x) = \boldsymbol{\eta}(x) + 2\mathbf{h}(x)\sqrt{8\pi G} \tag{4}$$

(donde G es la constante gravitacional expresada en unidades naturales, esto es, tales que $\hbar = c = 1$).[21]

Feynman demuestra sin dificultad que el tensor **g** tiene que satisfacer las ecuaciones de campo de Einstein. ¿Quiere esto decir que la aproximación clásica a la teoría cuántica de la gravedad esbozada por Feynman equivale a la relatividad general? Al parecer, Feynman pensaba que sí. Presenta la concepción geométrica original de la teoría de Einstein como una "interpretación" de la teoría "física" de campos que atribuye a sus venusianos.

> Como estos son en verdad dos aspectos de la misma teoría, podemos suponer que los científicos venusianos, después de desarrollar completamente su teoría del campo gravitacional, habrían descubierto el punto de vista geométrico. No podemos estar absolutamente seguros de ello, pues nunca se puede explicar el razonamiento inductivo; no se puede nunca explicar cómo hay que proceder, cuando uno solo sabe un poquito, para aprender aun más. En todo caso, el hecho es que el campo

> de spin 2 tiene este interpretación geométrica; esto no es algo que pueda entenderse sin más, es simplemente maravilloso.
>
> (Feynman *et al.* 1995, p. 113)

Feynman creía que "la interpretación geométrica no es realmente necesaria o esencial para la física" (ibid., p. 112). Pero en un respecto le atrae bastante. Nos recuerda que su punto de vista inicial "era que el espacio puede describirse como el espacio de la relatividad especial", esto es, como un espacio que, por alguna razón que me escapa, Feynman llama "galileano" (*Galilean*).

> En este espacio galileano puede haber campos gravitacionales $\eta_{\mu\nu}$ que tienen el efecto de que las reglas cambian de longitud y los relojes marchan más rápida o más lentamente. Por eso, al hablar de los resultados de experimentos estamos forzados a distinguir entre las escalas de las mediciones efectivas, las escalas físicas, y las escalas en que está escrita la teoría, las escalas galileanas. Ahora bien, la cosa es que son las coordenadas físicas las que tienen que reproducir siempre los mismos resultados. Puede ser conveniente que, al poner la teoría por escrito, supongamos inicialmente que las mediciones se llevan a cabo en un espacio que es galileano en principio; pero una vez que logramos predecir efectos reales, vemos que el espacio galileano no tiene ninguna significación. No tiene sentido que protestemos que la elección de coordenadas hecha por otro es insensata y absurda, solo porque no nos parece galileana. Pero si el otro insiste en considerarla galileana y le atribuye curvaturas a los campos, tiene perfecto derecho a ello, y es nuestro espacio el que le parece absurdo a él. Para cualquier resultado físico uno obtiene la misma respuesta, no importa qué combinaciones de números asigne uno inicialmente a las posiciones. Vemos, pues, que podría ser un adelanto filosófico (*it might be a philosophical improvement*) que lográsemos formular nuestra teoría desde un comienzo de tal modo que no haya un espacio galileano que entre en la especificación de la física; que tratemos siempre directamente con el espacio físico de las mediciones efectivas.
>
> (Feynman *et al.* 1995, p. 112)

Estas palabras de aprobación dedicadas a la interpretación geométrica de la gravedad no pueden ocultarnos las limitaciones que le impone el enfoque de Feynman. Si la solución **g** de las ecuaciones

de campo de Einstein tiene que satisfacer la condición (4), donde **η** es el tensor métrico del "espacio de la relatividad especial" y **h** es un campo tensorial sobre este espacio, la variedad subyacente a los tres campos **η**, **h** y **g** tiene que ser tal que admita la métrica **η**. Esta exigencia impone severas restricciones a la topología☆ de la variedad. Habitualmente, "el espacio de la relatividad especial" se considera homeomorfo☆ a $\mathbb{R}^4$ (esto es, al conjunto de todos los cuatruplos de números reales, provisto de la topología estándar generada por las esferas abiertas; véase §7.5, p. 163). Ello no es estrictamente necesario, pues es posible generar espaciotiempos planos de formas muy diferentes con solo identificar mutuamente diversos conjuntos de puntos, tal como, identificando los lados opuestos de un rectángulo, generamos la superficie de curvatura 0 de un cilindro o de la cinta de Moebius.[22] Pero su diversidad queda confinada dentro de límites muy estrechos, y no es nada comparable a la abigarrada multitud de los espaciotiempos topológicamente incompatibles que admite la relatividad general tal como fue concebida por Einstein. Así, los venusianos de Feynman podían concebir muy bien el espaciotiempo axiosimétrico de la primera solución exacta de las ecuaciones de campo de Einstein (Schwarzschild 1916*a*,*b*; Droste 1916) como un espaciotiempo de Minkowski del que se ha eliminado una recta temporaloide y que está encorvado por el campo **h** en las vecindades de esa recta. Pero tendrían que excluir, por topológicamente inadmisible, la solución exacta que inauguró la cosmología moderna (Einstein 1917*b*), pues la métrica plana **η** simplemente no puede definirse sobre la variedad homeomorfa a $\mathbb{R} \times \mathbf{S}^3$ requerida por esta solución. Tampoco podrían aceptar el modelo de un universo que sucesivamente se expande y se contrae propuesto independientemente por Friedmann (1922) y Lemaître (1927).[23] Por mucho tiempo, estos modelos de universo con *Big Bang* fueron más populares en la cosmología relativista que los modelos espacialmente abiertos y en expansión perpetua introducidos por Friedmann (1924) y por Einstein y de Sitter (1931), a pesar de que la astronomía observacional persistentemente ha fracasado en la empresa de registrar la presencia de la cantidad de materia que haría falta para que el universo entre en una etapa de contracción. Los físicos venusianos no sufrirían esta incertidumbre. Si un matemático venusiano

descubriese las soluciones de Friedmann de las ecuaciones de campo de Einstein, aquellas que se distinguen por sus espacios cerrados —topológicamente compactos— serían desechadas en el acto por ser físicamente imposibles, ya que no pueden estar realizadas en "el espacio de la relatividad especial", y solo las que presentan espacios abiertos —esto incluye los universos de Friedmann comúnmente llamados "abiertos" y "planos" (p. 124)— les parecerían viables. Además, en contraste con el descubrimiento de Friedmann, el de nuestro matemático venusiano no respondería a "consideraciones cosmológicas" motivadas física y filosóficamente como las de Einstein (1917*b*), pues en la Venus de Feynman un modelo cosmológico espacialmente finito pero ilimitado como el primero propuesto por Einstein ni siquiera entraría en consideración.

Tales especulaciones sobre desarrollos alternativos de la física en distintos planetas son, por cierto, en extremo aventuradas y quien tenga un mínimo sentido de la historia no dejará de mirarlas con sorna. Con todo, vale la pena recordar en el presente contexto que todas las corroboraciones más precisas de la relatividad general reunidas hasta la fecha se basan exclusivamente en aplicaciones de la solución de Schwarzschild (precesión de perihelios, deflección de la luz por la gravedad) o del principio de equivalencia (retardo gravitacional de los relojes). Los testimonios astronómicos que favorecen la cosmología del *Big Bang* —la recesión general de las galaxias (Hubble 1929), la radiación cosmológica de trasfondo (Penzias y Wilson 1964)— solo hablan a favor de ella si se los combina con la relatividad general, aceptada sobre la base de observaciones del sistema solar, entendidas a la luz de la solución de Schwarzschild. Sería pues aventurado decir que la cosmología moderna y sus exigencias de libertad topológica son suficientes para refutar el enfoque ageométrico de la teoría de Einstein patrocinado por Feynman. Por el contrario, no es claro que pueda justificarse la atribución de una determinada topología al universo en su conjunto invocando observaciones y experimentos emprendidos desde la tierra.

4

Modelos del mundo

Este capítulo se refiere a diversos espaciotiempos propuestos como modelos del universo por la cosmología relativista. El asunto se relaciona estrechamente con un tema capital de la filosofía de la ciencia, a saber, *el uso de modelos matemáticos para el estudio y representación de fenómenos naturales*. Sería un error, sin embargo, pensar que se trata de una subdivisión o especificación de este tema. Los espaciotiempos de que hablo aquí son por cierto una clase particular de modelos matemáticos y se los puede concebir sin dificultad como realizaciones específicas de un tipo de estructura matemática, esto es, como "modelos" en la acepción que tiene esta palabra en la llamada teoría de modelos. Pero el universo no es un fenómeno y, aunque adoptemos la opinión vulgar de que es la totalidad de los fenómenos, es de esperar que el tránsito a la totalidad cambie completamente las reglas del juego. Sin embargo, unas cuantas indicaciones sobre los modelos matemáticos de los fenómenos pueden arrojar alguna luz sobre los modelos espaciotemporales del universo.

Las estructuras matemáticas son conceptos perfectamente bien definidos cuya complejidad puede ser muy grande. El rasgo distintivo de las matemáticas no es la exactitud cuantitativa como creen los niños de escuela. Lo que hace única a las matemáticas entre todas las empresas intelectuales del hombre es la precisión de sus conceptos. Solo en el siglo XIX quedó esto bien claro para todos los filósofos, gracias a la formación de ramas nuevas de las matemáticas, como la geometría proyectiva y la topología, que no se ocupan con cantidades. Sin embargo, me atrevería a decir que ya Aristóteles pensaba en la perfecta nitidez conceptual de las matemáticas cuando distinguía tan enfáticamente entre predicados idóneos para la física, como σιμός ('ñato'), que solo se puede atribuir a narices, y predicados matemáticos como κοιλός ('cóncavo'), que es aplicable a cualquier superfi-

cie con la forma de una nariz ñata. "Difieren —dice Aristóteles— en que ñato es captado conjuntamente con la materia (pues lo que es ñato es una nariz cóncava), mientras que la cualidad de ser cóncavo es independiente de la materia perceptible. Entonces, si todas las cosas físicas se conciben como lo ñato —por ejemplo, nariz, ojo, rostro, carne, hueso y, en general, animal; hoja, raíz, corteza y, en general, planta (ninguna de las cuales admite una explicación que no involucre movimiento y, por ende, materia), es obvio cómo hay que buscar y definir el *qué es* de algo en el terreno de la física."[1] Ahora bien, la *materia* en el sentido de Aristóteles —ὕλη en griego— pertenece a todo lo que es susceptible de cambiar. La palabra se refiere a la potencialidad que tales cosas encierran de convertirse en otras que las que son; a su carácter impredecible y azaroso, a su inherente vaguedad y ambigüedad. Por lo tanto, cuando Aristóteles sostiene que todo concepto físico debe captar su objeto junto con la materia de su objeto está diciéndonos que los físicos deben dar cabida a la indefinición e indeterminación de las cosas que estudian y, por lo mismo, no pueden recurrir, salvo de paso, a los conceptos desmaterializados, cabalmente determinados de las matemáticas.

En el siglo XVI hubo un renacimiento de la filosofía aristotélica en muchas universidades europeas, católicas y protestantes. En el siglo XVII los fundadores de la física matemática moderna se alzaron en armas contra esa filosofía.[2] Como cristianos difícilmente podían aceptar que el ingrediente último de la creación fuese indeterminado y potencial. Como criatura *actual* de la voluntad de Dios, la materia cristiana tiene que ser portadora de todas las cualidades necesarias para cumplir el propósito divino, y de ninguna otra. La deidad cristiana sabía seguramente lo que quería cuando creó el mundo y tenía todo el poder necesario para producir un material exactamente adecuado a sus fines. La oposición a Aristóteles es evidente en el dicho de Kepler de que "Dios es el eterno geómetra" (Ὁ θεὸς ἀεὶ γεωμητρίζει) y en la aseveración de Galileo de que "el libro de la naturaleza" está escrito "en lenguaje matemático" y "sus caracteres son triángulos, círculos y otras figuras geométricas, sin las cuales sería humanamente imposible entender de él ni una sola palabra" (Galileo 1623, § 6; EN, VI, 232). Los físicos del siglo XVII pensaban sin duda, en sus momentos

de mayor optimismo, que eventualmente descifrarían la escritura de Dios. Su esperanza estaba avalada por la creencia de que Él "emplea habitualmente los medios más primordiales, más simples y más fáciles" (Galileo, EN, VIII, 197). Comprendían, sin embargo, que como criaturas de inteligencia mediocre y experiencia muy estrecha, tenían que avanzar lentamente y con paso suave en su empresa científica. En contraste con Aristóteles y sus seguidores, no se adelantarían a tratar con los fenómenos naturales en la forma completa y concreta en que normalmente se nos dan, sino que en vez de eso intentarían aislar los procesos más simples cuya combinación genera los muy complejos que vemos y tratarían de entender a cada uno por sí mismo, para luego, a la larga, mostrar cómo el espectáculo abigarrado de la naturaleza es la resultante de todos.

Un buen ejemplo de este enfoque ofrece la teoría de la caída libre de Galileo, expuesta por su portavoz Salviati en la tercera jornada de los *Discorsi* (EN, VIII, 190ss.). Salviati lee una definición formal de movimiento uniformemente acelerado y deriva de ella, combinada con verdades familiares de la aritmética y la geometría, una relación funcional bien determinada entre la duración de un movimiento de esta clase y la distancia recorrida y la velocidad alcanzada por el móvil. La teoría asevera que este concepto de movimiento uniformemente acelerado se aplica cerca de la superificie de la tierra a todo cuerpo que *solamente gravita* y no está sujeto a otras influencias como la resistencia del aire o el roce del suelo sobre el que rueda o se desliza. Obviamente, estas influencias no pueden suprimirse completamente en un caso particular dado, pero un experimentador ingenioso se las arregla para reducirlas. Supongamos pues que, cuando el roce del suelo y la resistencia del aire disminuyen, los tiempos, distancias y velocidades efectivamente medidos están cada vez más cerca de satisfacer la relación funcional prescrita por el concepto matemático. En tal caso, uno concluirá que este representa adecuadamente la ley de la caída libre cerca de la superficie de la tierra, aunque no haya un solo cuerpo que efectivamente caiga en esa vecindad en exacta concordancia con él.

Este método galileano de concebir matemáticamente los fenómenos naturales simplificados, depurados, aislados y —hay que decir-

lo— abstractos fue continuado y perfeccionado por sus grandes sucesores, Torricelli, Pascal, Huygens y, sobre todo, por Newton. De hecho, la contribución de este último fue tan grande y tuvo tanta influencia que el historiador I. B. Cohen atribuye el método a Newton más bien que a Galileo.[3] Tal como los geómetras investigan las propiedades de toda clase de figuras no importa que existan o no en la naturaleza, así Newton llevó adelante sus análisis de diversas configuraciones de masas y fuerzas en el espacio y en el tiempo sin prestarle la menor atención al hecho de que tal vez no existen. La selección de las configuraciones estudiadas se guiaba por el propósito de hallar una o más que proporcionasen una representación matemática suficientemente adecuada de los movimientos del sistema solar; pero también estuvo constreñida por la necesidad de entender las matemáticas de unas cuantas situaciones irreales simples para resolver las más realistas y complejas. Estas últimas, claro, solo calzan con los fenómenos dentro de algún margen pragmáticamente aceptable de error. Otro tanto podría quizás decirse de las otras, si se permite un margen de error lo bastante grande. Por ejemplo, no se hallará en nuestro entorno un cuerpo sobre el que actúe una fuerza dirigida a un centro infinitamente distante; sin embargo, el movimiento de las balas de cañon de un punto a otro de un campo de batalla y el movimiento durante un breve lapso de tiempo del sistema formado por la tierra y la luna pueden representarse tolerablemente bien mediante configuraciones sometidas a esa hipótesis patentemente ficticia (Newton 1726, pp. 94, 117).

La eficacia del método de Galileo y Newton ha sido confirmada por los éxitos de la física experimental y sus aplicaciones técnicas. Pero sus creadores evidentemente no pretendían abrazar la totalidad por esta vía; antes bien, postergaron y tal vez renunciaron a la meta aristotélica de entender el mundo natural en su integridad concreta.[4] Newton lo vio claramente. Consideraba a la gravedad como una fuerza natural entre muchas otras, aún por desentrañar; y mientras el poeta Pope veía en él la luz creada por Dios para iluminar la oscuridad de la naturaleza, Newton se percibía a sí mismo como un niño pequeño que jugaba con guijarros y conchillas en la playa del enorme y desconocido "océano de la verdad".[5] Por lo demás, contra todo lo que

pudiéramos pensar sobre la vocación cosmológica inherente en una teoría de la gravedad, la ciencia newtoniana, en sus dos siglos de predominio, prestó muy poca atención a la estructura global del universo. De hecho, sus cultivadores apenas se sintieron afectados por la paradoja que parece inevitable en una cosmología newtoniana, a saber, que si las estrellas están todas confinadas dentro de un espacio finito tienen que caer unas sobre otras, formando una sola bola densísima, mientras que, si están distribuidas al azar sobre un espacio infinito, el firmamento entero ha de brillar como el sol. Tampoco hay indicios de que algún físico de primera línea del siglo XIX, por ejemplo, Fourier o Ampère, Faraday o Helmholtz, Kelvin o Maxwell, haya protestado contra la conclusión de Kant de que la cosmología es una pseudociencia intrínsecamente contradictoria, impulsada por una ilusión de la razón.

Esta indiferencia de los fisicos a la cosmología terminó abruptamente el 8 de febrero de 1917, cuando Einstein presentó a la Academia Prusiana de Ciencias sus "Consideraciones cosmológicas sobre la teoría general de la relatividad". Como todos los grandes artículos de Einstein, es tremendamente innovador. También es maravillosamente claro y —al igual que los otros— constituye un ejemplo brillante de modelado matemático conforme al método de Galileo y Newton. Pero el objeto al cual el método se aplica, el sistema físico aislado que, tras una simplificación idónea, se representa aquí mediante un modelo sumamente idealizado, no es otro que el universo entero. La totalidad de la materia se identifica con un fluido sin presión que llena homogéneamente una ☆variedad riemanniana cuatridimensional☆ cuya métrica concuerda localmente en todas partes, hasta las cantidades de primer orden, con la del espaciotiempo minkowskiano de la relatividad especial. Pero este espaciotiempo no es homeomorfo con $\mathbb{R}^4$ como el de Minkowski —y el de Newton— sino con $\mathbb{R} \times \mathbf{S}^3$. El fluido cósmico describe una congruencia☆ de geodésicas temporaloides, cualesquiera dos de las cuales se mantienen para siempre a una distancia recíproca fija. El haz de geodésicas de la congruencia está seccionado ortogonalmente por subvariedades tridimensionales isométricas que no se cortan entre sí. Cada una de estas secciones o "tajadas" espacialoides puede verse como una representación del

espacio cósmico en un instante determinado. La gran novedad es que las tajadas son homeomorfas a $\mathbf{S}^3$. Así, en cualquier instante, el universo de Einstein (1917*b*) ocupa un volumen finito pero en ningún lugar tropieza con un límite.[6]

Ningún físico había propuesto antes un modelo espaciotemporal del universo y nos gustaría naturalmente saber qué indujo a Einstein a presentarlo. ¿Por qué esta súbita preocupación con la estructura global del espaciotiempo? ¿La motivan acaso nuevos resultados experimentales? Y, si no es asi, ¿qué provocó el inesperado arranque de Einstein de abrazar la totalidad? Las consideraciones cosmológicas de Einstein pertenecen a su teoría de la gravedad, más conocida como relatividad general. La métrica riemanniana de su modelo espaciotemporal del universo es una solución de las ecuaciones de campo de esta teoría. Einstein no contaba con ningún testimonio favorable a esta particular solución; no tenía indicios, por ejemplo, de que el universo sea homeomorfo☆ a $\mathbb{R} \times \mathbf{S}^3$. Por lo tanto, todo el respaldo empírico de su modelo espaciotemporal consistía en la evidencia que Einstein podía reunir a favor de la relatividad general. Ahora bien, en 1917 había buenos resultados de laboratorio que corroboraban la relatividad especial y, por tanto, también cualquier teoría que, como la relatividad general, concordara con la relatividad especial en un pequeño entorno de cada evento; y había también una evidencia avasalladora a favor de la teoría de la gravedad de Newton y, por tanto, también de cualquier teoría que, como la relatividad general, concuerda bien con la teoría de Newton en las circunstancias familiares en que dicha evidencia se había recogido. Pero la justificación empírica de la relatividad general en cuanto ella discrepa de esas teorías anteriores era a la sazón muy magra. Reposaba únicamente sobre la anomalía del perihelio de Mercurio, que la teoría de Newton era incapaz de explicar pero que se infiere de la de Einstein. El avance anómalo del perihelio de ese planeta suma 43 segundos de arco por siglo, menos del 1% de la precesión total observada. Para la astronomía esta pequeña diferencia no es desdeñable, pero como testimonio de la figura del espacio cósmico es irrisoria. Además, la predicción relativista de la precesión del perihelio de Mercurio se basa en una solución de las

ecuaciones de campo de Einstein (ECE) en su versión original, comunicada a la academia de Berlín el 25 de noviembre de 1915:[7]

$$R_{ij} = -\kappa\left(T_{ij} - \tfrac{1}{2} g_{ij} T_r^r\right) \qquad (0 \leq i, j \leq 3) \tag{1}$$

Pero la métrica del primer modelo cosmológico de Einstein es una solución de la versión corregida de las ecuaciones (1) que él introdujo en el artículo de 1917, justamente para que fuese posible obtener esa solución:

$$R_{ij} - \lambda g_{ij} = -\kappa\left(T_{ij} - \tfrac{1}{2} g_{ij} T_r^r\right) \tag{2}$$

La nueva versión equivale a las ecuaciones

$$R_{ij} - \tfrac{1}{2} g_{ij} R + \lambda g_{ij} = -\kappa T_{ij} \tag{2′}$$

en las cuales el lado izquierdo exhibe la forma más general que puede tomar un tensor simétrico de rango 2 dependiente de la métrica y cuya ☆derivada covariante☆ sea idéntica a 0, siendo este último un requisito que el tensor del lado derecho tiene que satisfacer si representa la distribución de la materia y la energía. λ es la llamada "constante cosmológica". Si ponemos $\lambda = 0$, recuperamos las ecuaciones originales. Por lo tanto, la corrección consistió en poner $\lambda \neq 0$. Por otra parte, la órbita observada de Mercurio solo puede obtenerse de las ecuaciones (2) si el valor de λ es extraordinariamente pequeño. En efecto, como Einstein seguramente sabía, las observaciones del sistema solar imponen a λ una cota superior tan baja que no dan ni el más mínimo motivo para suponer que λ difiere de 0.[8]

No parece pues que Einstein haya dirigido su pensamiento hacia la cosmología urgido por la presión de los hechos. Ni Einstein sugirió nunca que así fuese. Cuando escribió las "Consideraciones cosmológicas" se conocía solo una solución exacta de las ecuaciones (1). Estaba calculada para representar un sistema gravitacional aislado formado por un cuerpo muy grande y otro de un tamaño insignificante. Schwarzschild (1916*a*, 1916*b*) encontró esta solución postulando una métrica del espaciotiempo que

(*a*) fuera esféricamente simétrica en el espacio,
(*b*) no variara con el tiempo, y
(*c*) convergiera a la métrica plana de Minkowski en el infinito espacial.[9]

Einstein juzgó objetable la condición (*c*). Que el espaciotiempo sea plano en el infinito espacial significa que una partícula de prueba se vería allí constreñida por la métrica a trazar una cosmolínea recta aunque el resto de la materia diste infinitamente de ella. Según Einstein, esto querría decir que la estructura del espaciotiempo es capaz de guiar el movimiento de la materia sin estar a su vez determinada por la distribución de esta. Ello constituiría una violación crasa de una tesis que Einstein atribuía a Mach y consideraba como uno de los principios fundamentales de la relatividad general. Este "principio de Mach" (*Mach'sches Prinzip*) puede enunciarse así:

> El campo [métrico] está determinado *exhaustivamente* por las masas de los cuerpos.
>
> (Einstein 1918*e*, p. 241)

G. D. Birkhoff (1923) demostró que la condición (*a*) de la solución de Schwarzschild implica las condiciones (*b*) y (*c*). Ello entraña, obviamente, que si el campo esféricamente simétrico de Schwarzschild está determinado por la materia concentrada supuestamente en el centro de simetría, esa materia también determina la convergencia del campo a una métrica plana en el infinito espacial. Si esto es así, no puede decirse que la condición (*c*) esté en conflicto con el principio de Mach, según Einstein lo entiende. Afortunadamente, en 1917 el teorema de Birkhoff no se conocía, y Einstein pudo todavía pensar que la relatividad general realmente afrontaba aquí una dificultad, para resolver la cual inventó la cosmología moderna.

Einstein expresa así su rechazo de un espaciotiempo plano en el infinito espacial: "En una teoría de la relatividad que sea consistente consigo misma no puede haber inercia *con respecto al* '*espacio*', sino solo inercia de las masas *respecto las unas de las otras*. Por lo tanto, si transporto una masa a una distancia suficientemente grande de todas las demás, su inercia debe caer a cero" (1917*b*, p. 145). Si el campo

gravitacional es isotrópico en el espacio en torno a cada punto, hay un sistema de coordenadas $x = (x^0,x^1,x^2,x^3)$, relativamente al cual la métrica espaciotemporal está dada por el ☆elemento de línea☆:

$$ds^2 = B(dx^0)^2 - A((dx^1)^2 + (dx^2)^2 + (dx^3)^2) \qquad (3)$$

Einstein muestra que, en tal caso, la masa inercial de una partícula cae a cero en el infinito espacial si y solo sí A también se anula allí, mientras que B crece por sobre todo límite asignable. En colaboración con J. Grommer, Einstein emprendió un estudio de los campos estáticos esféricamente simétricos cuya métrica pudiera degenerar de este modo en el infinito espacial. Concluyó que "tales condiciones de frontera son absolutamente inadmisibles para el sistema de las estrellas fijas" (1917*b*, p. 146). En rigor, como sabemos ahora, las condiciones $A \to 0$ y $B \to \infty$ no están excluidas solo en virtud de alguna propiedad observable de las estrellas fijas, sino que, por el teorema de Birkhoff, son *lógicamente incompatibles* con la simetría postulada.

Se le ocurrió entonces a Einstein que una métrica riemanniana puede estar realizada en una variedad espacialmente compacta☆. En un espaciotiempo con esta propiedad no hay un infinito espacial y por tanto no hay que postular condiciones de frontera que se cumplan allí. Einstein se propuso hallar una solución de sus ecuaciones del campo gravitacional que cumpliese las siguientes condiciones:

(i) El campo métrico está definido sobre una variedad que admite una foliación[10] en hipersuperficies espacialoides que llamaré 'tajadas canónicas'. Cada tajada canónica es una ☆variedad topológica☆ tridimensional compacta☆.
(ii) El campo métrico es homogéneo e isotrópico sobre cada tajada canónica o, como también podemos decir, es *espacialmente isotrópico y homogéneo*.[11]
(iii) El campo métrico no varía de una tajada canónica a otra.

Con la condición (i) se busca, claro está, obviar la necesidad de prescribir condiciones de frontera en el infinito espacial. Si la materia y la radiación está distribuidas isotrópicamente en torno nuestro, la condición (ii) se deduce de este hecho combinado con el principio

"copernicano"[12] según el cual no ocupamos una posición privilegiada en el universo. La condición (iii) refleja de una manera simplificada e idealizada un supuesto hecho que Einstein da por sentado, a saber, que las estrellas se mueven al azar en todas direcciones a velocidades relativas muy inferiores a la velocidad de la luz. Hoy percibimos un conflicto entre las observaciones que respaldan este supuesto y las que pueden aducirse en favor de la condición (ii). La astronomía extragaláctica actual enseña en efecto que la materia del universo a gran escala está isotrópicamente distribuida en torno nuestro, pero rechaza terminantemente la pretensión de que las estrellas se mueven con velocidades bajas y distribuidas al azar. Según la llamada "ley de Hubble", las galaxias en general se alejan de nosotros a velocidades que aumentan indefinidamente con la distancia. Las bajas velocidades compatibles con la condición (iii) se observan empero dentro de nuestra galaxia, que era la fuente de toda la información astronómica utilizada por Einstein en 1917. Pero las observaciones confinadas a nuestra propia galaxia no permiten inferir una distribución isotrópica de la materia en torno nuestro, pues la vasta mayoría de sus estrellas se encuentra localizada en una faja más bien angosta del firmamento. Por eso pienso que en el artículo de Einstein la condición (ii) no responde a lo que la experiencia le sugería sino al deseo de habérselas con un problema matemáticamente manejable.[13] Para satisfacer las tres condiciones a la vez Einstein tuvo que reemplazar las ECE (1) con las ECE (2). Observa, sin embargo, que las ecuaciones (1) son compatibles con la condición (i) y que el término nuevo λg_{ij} (con $\lambda \neq 0$) se requiere solo para cumplir la condición (iii). J. B. Barbour (1990, p. 63 n. 7) señala acertadamente que esta sugestiva observación de Einstein bien puede haber estimulado a Alexander Friedmann a hallar su familia de soluciones de las ecuaciones (2) para un fluido homogéneo e isotrópico sin presión. Yo iría más lejos: Einstein difícilmente habría publicado esa observación si no hubiese hallado una solución particular de las ecuaciones (1) que satisfaga las (i) y (ii), pero no la condición (iii).

Las soluciones de Friedmann a las ECE (2) aparecieron en dos artículos en la revista *Zeitschrift für Physik* (Friedmann 1922, 1924).[14] Friedmann demuestra que la única solución que cumple con las con-

diciones (i), (ii) y (iii) es la de Einstein (1917*b*).[15] La solución de Einstein presupone que $\lambda > 0$. Si descartamos la condición (iii) hay una infinidad de soluciones con λ arbitrario, que comparten la propiedad siguiente: las tajadas canónicas son hipersuperficies espacialoides de curvatura positiva constante que varía continuamente a lo largo de las cosmolíneas geodésicas de la materia.[16] Evidentemente, el volumen de las tajadas canónicas crecerá según decrecen sus respectivas curvaturas, y viceversa.[17] En su segundo artículo, Friedmann descarta la exigencia de compacidad incluida en la condición (i). Obtiene entonces una infinidad de soluciones con λ arbitrario, en que las tajadas canónicas tienen curvatura constante negativa que varía continuamente a lo largo de las cosmolíneas geodésicas de la materia. La segunda familia de soluciones de Friedmann puede extenderse de un modo natural para que incluya espaciotiempos cuyas tajadas canónicas tienen curvatura constante igual a cero, lo que hoy suelen llamarse "universos planos" (Robertson 1929). El tipo de evolución que tales universos pueden exhibir y su dependencia de λ son iguales, ya sea que la curvatura en cuestión sea igual o menor que 0 (Robertson 1933, p. 75). En ambos casos, las tajadas canónicas se extienden hasta el infinito y por tanto no pueden propiamente aumentar o disminuir de volumen. Sin embargo, hay un sentido en que es lícito decir que los modelos cosmológicos de esta familia se expanden o se contraen: en cualquiera de ellos, las cosmolíneas geodésicas de la materia no son equidistantes por pares como en el universo estático de Einstein, sino que, cada vez que cortan una tajada canónica dada, o bien todas se acercan mutuamente, o bien van todas alejándose. La existencia de esta familia de soluciones exactas de las ecuaciones (1) y (2) prueba, según Friedmann, que, a menos que agreguemos algunos supuestos suplementarios, las ECE no nos permiten llegar a conclusión alguna sobre la finitud o infinitud del mundo en que vivimos (1924, p. 332). Ambas familias de soluciones de Friedmann parten de la hipótesis de que la materia consiste en polvo incoherente (sin presión); pero Lemaître (1927) las derivó de nuevo en una versión más general que abarca el caso de que la materia sea un fluido perfecto.[18]

A diferencia de las "Consideraciones cosmológicas" de Einstein, los artículos de Friedmann no combinan en un delicado equilibrio el razonamiento matemático, los datos empíricos y las aspiraciones filosóficas. Más bien parecen ejercicios matemáticos de solución de ecuaciones diferenciales. Sin embargo, sus implicaciones filosóficas son bastante claras. Aun si sometemos el tensor de energía T_{ij} a una drástica idealización simplificadora, las ECE (1) y (2) admiten una enorme variedad de soluciones cosmológicas que ni siquiera comparten todas la misma topología global. Más tarde nos hemos enterado de muchas alternativas extremas que corresponden a supuestos menos simplistas; pero los artículos de Friedmann fueron prueba suficiente de que la incursión de Einstein en la cosmología dejaba mucho sin decidir. Las soluciones de Friedmann fueron ignoradas, a pesar de publicarse en la misma prestigiosa revista en que, pocos años más tarde, Heisenberg y sus amigos de Göttingen sentarían las bases de la mecánica cuántica. Algunos científicos probablemente dieron fe a la descomedida afirmación de Einstein —en la misma revista— de que la obra de Friedmann reposaba en un error matemático,[19] pues siguieron prestando atención exclusivamente a los modelos cosmológicos de Einstein y de Sitter (véase Ellis 1989, pp. 379–380). Pero la gran mayoría debe haber pensado que la cosmología relativista no valía la pena. No era fácil que pensasen otra cosa mientras la relatividad general no tuviese más respaldo empírico que el de los tests clásicos de la solución de Schwarzschild y el principio de equivalencia. Porque no importa cuán fuertemente los corrobore la experiencia, nos dicen bien poco sobre la estructura global del universo.

En 1927, Georges Lemaître, que no conocía las soluciones de Friedmann, las redescubrió por su cuenta y las publicó nuevamente en los Anales de la Sociedad Científica de Bruselas. Lemaître reaccionaba a un cambio radical ocurrido en la astronomía. Gracias sobre todo a Edwin Hubble, la vieja doctrina de Kant y Herschel según la cual la mayoría de las nebulosas observables en el cielo son en realidad "universos-islas", como nuestra Vía Láctea, había logrado finalmente convencer a casi todos los astrónomos.[20] Hubble sostuvo asimismo que todos esos otros sistemas estelares o "galaxias" —como se dio en llamarlos, utilizando en las lenguas modernas como sustan-

tivo común el nombre propio del nuestro en griego[21]— están distribuidos isotrópicamente en torno a nosotros; del corrimiento sistemático hacia el rojo del espectro de la luz que recibimos de ellos infirió que se alejan de nosotros con velocidades proporcionales a la respectiva distancia. Ahora por fin la cosmología relativista podía apoyar en observaciones genuinas el postulado de que la materia está distribuida isotrópicamente y, por tanto, en virtud del principio copernicano, también homogéneamente. Por otra parte, las soluciones de Friedmann y Lemaître ofrecían una explicación simple de la recesión de las galaxias: si describen cosmolíneas geodésicas en un espaciotiempo de Friedmann-Lemaître en expansión, todas huirán de cada una tanto más rápidamente cuanto más alejadas estén de ella. De este modo la relatividad general hacía posible entender la desconcertante fuga de las galaxias como una consecuencia necesaria de la misma geometrodinámica que fija su trayectoria a cada cuerpo en caída libre. Sin embargo, la obra de Lemaître pasó en buena medida desapercibida hasta 1930, cuando Eddington y McVittie estudiaron la estabilidad de la solución cosmológica de Einstein (1917*b*). Antes de completar su investigación se percataron de la existencia "de un artículo del Abate G. Lemaître que presenta una solución notablemente completa de los varios problemas relacionados con las cosmogonías de Einstein y de Sitter; aunque no lo dice expresamente, a la luz de sus fórmulas es obvio que el universo de Einstein es inestable" (Eddington 1930, p. 668). Poco más tarde, Einstein y de Sitter (1932) publicaron conjuntamente un artículo que recomienda la solución de Friedmann (sin presión) con $\lambda = 0$ y tajadas canónicas planas.

Ello no obstante, pasaron cuarenta años antes de que las soluciones de Friedmann y Lemaître se convirtiesen en *el modelo estándar* del universo. En su brillante panorama de la cosmología relativista, Robertson modestamente reclama solo esto en su favor: "Espero mostrar que, [con ayuda de los resultados de Hubble y] guiados por unas pocas suposiciones y extrapolaciones que parecen naturales, podemos llegar a un sistema intrínsecamente razonable de cosmología relativista que no está en serio conflicto con la astrofísica moderna" (1933, p. 62). Pero otro tanto podía reclamarse en pro de las teorías cosmológicas propuestas en los años treinta por Milne, Dirac, Jordan

y otros; y también por cierto de la controvertible teoría del "estado estacionario" (*steady state*) publicada independientemente después de la Segunda Guerra Mundial por Bondi y Gold (1948) y por Fred Hoyle (1948), la cual tuvo bastante aceptación en el Reino Unido hasta mediados de los años sesenta. La vigorosa aprobación pública de la cosmología relativista desde 1970 se debe en parte al respaldo que le han dado a la relatividad general los nuevos experimentos practicados en la tierra y la observación más precisa del sistema solar; en parte también, pienso, a la diseminación de unos cuantos libros de textos excelentes, en particular los de Steven Weinberg (1972) y de Misner, Thorne y Wheeler (1973); pero sobre todo al hecho —descubierto por Penzias y Wilson (1965) y bellísimamente confirmado por mediciones efectuadas fuera de la atmósfera terrestre— de que vivimos sumergidos en un océano de radiación termal sumamente isotrópica de muy baja temperatura. Alpher y Herman (1948) previeron este fenómeno en el curso de sus investigaciones —inspiradas por Gamow— sobre el origen de los elementos más pesados que el hidrógeno; pero su predicción quedó enterrada en una revista científica.[22] Fue predicho nuevamente por los físicos del grupo de Robert Dicke en Princeton, quienes hacia 1965 revivieron las ideas de Gamow sobre la formación del helio, etc. por nucleosíntesis en un crisol cósmico muy denso y muy caliente.[23] Por eso, cuando Penzias y Wilson registraron estupefactos la presencia de radiación fría procedente de todas direcciones con la misma intensidad, la teoría física ya tenía disponible una explicación, suministrada conjuntamente por la física cuántica del núcleo atómico y la relatividad general. Merced a ello, su descubrimiento les valió el premio Nobel (otorgado a Penzias y Wilson en 1978).

En el último cuarto de siglo el modelo estándar ha tenido un éxito enorme —científico, periodístico y literario—, encantando a las clases educadas no menos, pienso, que el poema de Gilgamesh, el Génesis y la *Teogonía* de Hesíodo deleitaron a sus contrapartes en la antigüedad. Por supuesto, no todos están convencidos. Unos pocos astrónomos rechazan la recesión de las galaxias, otros dudan del origen cósmico de la radiación térmica de fondo. Desde un punto de vista filosófico debemos atender especialmente a la opinión expresada con

vigor por Massimo Pauri de que "el universo *como un todo* no puede considerarse como un *objeto científico* en ningún sentido que estas palabras hayan tenido en el curso del desarrollo histórico de la física" (1991, p. 291). No sin resistencia, he llegado paulatinamente a creer que Pauri tiene toda la razón. Al menos, estoy convencido de que, aunque la cosmología relativista gana fácilmente cualquier concurso de belleza entre las teorías físicas que hay, ella obtendrá una nota final mucho más baja si también se dan puntos por el poder de convicción y la confiabilidad dentro de un margen de error sensato y útil. Esta no es más que una corazonada mía, inspirada quizás por mis ideas previas sobre la naturaleza del modelado matemático y los tests experimentales. Sin embargo, expondré para terminar algunos motivos de mi escepticismo.

Si he narrado bien la historia, lo que decidió la partida en favor del modelo estándar fue el descubrimiento de la radiación isotrópica de fondo por Penzias y Wilson. Este constituye un testimonio favorable a la cosmología relativista solo si se lo entiende a la luz de la física cuántica, como radiación térmica ajustada a la ley de Planck y que es un vestigio de las reacciones nucleares que generaron la abundancia actual de elementos livianos más pesados que el hidrógeno. Pero la fisica cuántica es lógicamente incompatible con la relatividad general. La inferencia científica que conduce al modelo estándar utiliza, pues, premisas que son incompatibles entre sí. En la física ordinaria, que se ocupa de aspectos particulares de los fenómenos dentro de márgenes admisibles de error, uno puede sin duda jugar con distintos modelos extraídos de teorías mutuamente contradictorias pero que se dejan combinar bien con la aproximación aceptada. Sin embargo, dudo de que este procedimiento pueda producir resultados creíbles acerca de la totalidad de las cosas. Dado el éxito de la física cuántica en nuestros laboratorios de alta energía nadie espera que la relatividad general sea válida en el universo muy temprano, muy caliente y muy denso. Por otra parte, no está claro cuán denso y cuán caliente tiene que estar el universo para que las ECE fallen en una medida significativa.[24] Y, por supuesto, sin tests no podemos anticipar qué parte de la geometría del universo vaya a sobrevivir cuando las ECE se desmoronen.

Dejando de lado, empero, el problema de la coexistencia de la relatividad general con el valor finito de la constante de Planck, consideremos la cosmología relativista en sus propios términos, como aplicación de la relatividad general. La aceptación de esta teoría descansa sobre el cumplimiento satisfactorio de predicciones calculadas en el contexto de la solución de Schwarzschild (con el tensor de Ricci☆ igual a 0) y de la aproximación parametrizada post-newtoniana al sistema solar (con los parámetros pertinentes fijados conforme a la relatividad general).[25] Pero el fluido homogéneo e isotrópico del modelo estándar no tiene cabida para sistemas solares, mucho menos para campos de Schwarzschild. McVittie (1933) publicó una solución de las ECE que superpone un espaciotiempo de Friedmann y el campo de Schwarzschild de un punto-masa único.[26] Sin embargo, tal esquema tiene muy poco que ver con el mundo real. Una idealización más realista trataría el universo como una suerte de mosaico formado con trocitos que reproducen la región central del campo de Schwarzschild, a razón de uno para cada estrella. Lindquist y Wheeler (1957) construyeron un espaciotiempo en que los espacios constan de 120 "celdas" de este tipo, cada una de las cuales tiene toda su masa concentrada en un agujero negro central, mientras el resto de la celda está vacía. Esta construcción simula satisfactoriamente un espaciotiempo cerrado (espacialmente compacto) de Friedmann.[27] Pero nuestro universo, claro está, es bastante más complicado que esto. Por un lado, no todas las estrellas están tan aisladas como el Sol. ¿Puede modelarse satisfactoriamente el campo de una estrella doble recurriendo a la solución de Schwarzschild? Por otro lado, las estrellas están agrupadas en galaxias, que forman cúmulos de galaxias, que a su vez se agrupan en supercúmulos, de modo que las galaxias en general están situadas en la vecindad de superficies separadas por inmensos espacios vacíos. En el universo homogéneo del modelo estándar, cada cosmolínea geodésica representa un supercúmulo de galaxias. ¿Sería justo construir un mosaico semi-friedmanniano en que cada supercúmulo esté representado por un campo de Schwarzschild? ¿O cabe esperar más bien que la compleja articulación de estos juegue un papel significativo en la construcción de la geometría global?

En los últimos treinta años los cosmólogos han tendido a tratar el problema que suscita la acumulación ostensible de la materia en un sentido exactamente contrario al que acabo de esbozar. En vez de calcular cómo pueden combinarse los cúmulos para generar la geometría de un fluido homogéneo e isotrópico, han dado ese fluido por supuesto y han tratado de explicar la formación de los cúmulos. Por cierto, si se pudiera mostrar que la evolución del modelo estándar implica la aglomeración del fluido homogéneo en estrellas, galaxias y cúmulos y supercúmulos de galaxias, se disiparían las dudas que he expresado. Si los cúmulos locales son el fruto del desarrollo geometrodinámico de un universo de Friedmann, este podrá ciertamente recuperarse combinando todos los cúmulos en un mosaico gigantesco. Sin embargo, no parece que las explicaciones propuestas para la formación de las galaxias basten para sostener esa conclusión.

Sea de ello lo que fuere, la contribución más publicitada a la cosmología actual quiere resolver otro problema en cierto modo opuesto al que hemos estado considerando. En cada momento de la evolución del modelo estándar, algunas partes de la materia y la radiación no han tenido tiempo aún para interactuar. Por ejemplo, las dos galaxias más lejanas que hoy logremos divisar apenas en lados opuestos del firmamento no pueden observarse aún la una desde la otra. ¿Cómo llegan entonces a ser tan parecidas? En particular ¿cómo pudo la radiación cósmica de fondo en esas regiones no interactuantes del universo alcanzar el estado de equilibrio térmico que le atribuimos? Alan Guth (1982) propuso una solución inesperada a este problema —y varios otros— suponiendo audazmente que la ecuaciones de campo de Einstein operan en un universo regido por una teoría cuántica de gran unificación.[28] Conforme a los supuestos de Guth, el universo se expandió exponencialmente durante un período "inflacionario" brevísimo (menos de 10^{-30} s), trascurrido el cual la expansión se ajusta al modelo estándar. Antes de que esto ocurra, aun las partes más distantes del universo observable hoy estaban lo bastante cerca como para entremezclarse mediante los familiares procesos térmicos. El modelo original de Guth generaba empero grandes inhomogeneidades, que lo hacían inadmisible (Blau y Guth 1987, p. 550). Para superar esta dificultad se propusieron otros modelos inflacionarios,

el más notable de los cuales es quizás uno debido al físico ruso Andrzej Linde. En este, "la estructura *local* del universo está determinada por la inflación" gobernada por las ECE; pero "su estructura *global* está determinada por efectos *cuánticos*". Tal como lo ve Linde, la evolución del universo no tiene fin y puede no tener un comienzo. "Como resultado, el universo se divide en muchos dominios diferentes (miniuniversos) de tamaño exponencialmente grande, dentro de los cuales están realizados todos los estados (metaestables) posibles del vacío" (Linde 1987, p. 607; véase también Linde 1983, 1986). Tales miniuniversos pueden poseer propiedades físicas básicas muy diversas, por ejemplo, diferir en el número de dimensiones. Uno de ellos se convierte a través de la inflación en el universo cuatridimensional, más o menos de tipo Friedmann-Lemaître, en que vivimos, el cual, en rudo contraste con sus hermanos, exhibe la combinación muy improbable de cualidades que hace posible la vida humana. La especulación de Linde es fascinante desde un punto de vista estético y religioso. Pero arroja un balde de agua fría sobre las ambiciones de la cosmología científica. De atenernos a ella, la investigación empírica no puede extenderse más allá de nuestra minifracción del universo. Si la visión de Linde es en absoluto plausible, el alcance de la ciencia se habría ampliado muchísimo desde el tiempo, no hace todavía dos siglos, en que Augusto Comte confinaba la astronomía a nuestro sistema planetario; pero aún así estaría impedida para siempre de abarcar el universo entero.

5

Los universos de Friedmann: derivando la geometría

5.1 El modelo estándar de la cosmología actual

El conocido libro de texto de P. J. E. Peebles (1993) gira en torno al modelo estándar cuyos cuatro *elementos principales* el autor resume así:

> **[1]** A gran escala, la distribución media de las masas es prácticamente homogénea.
> **[2]** El universo se está expandiendo; mejor dicho, la distancia media L entre las partículas conservadas aumenta con el tiempo a la velocidad $dL/dt = HL$. El factor de proporcionalidad H depende del tiempo y su valor actual es igual a la llamada constante de Hubble H_0.
> **[3]** La dinámica del universo en expansión está descrita por la teoría general de la relatividad de Einstein.
> **[4]** El universo se ha expandido desde un estado caliente y denso en que su masa estaba dominada por radiación térmica de cuerpo negro.
>
> (Peebles 1993, pp. 5–6)

El modelo estándar está respaldado por tres tipos de *fenómenos* (cf. Newton 1726, pp. 390ss.). Naturalmente, para que puedan darles respaldo, cada fenómeno tiene que entenderse a la luz de uno o más supuestos teóricos. Así, puede decirse que el cuerpo del modelo estándar, compuesto por los cuatro elementos nombrados, descansa en tres soportes o muletas empíricas, cada una de las cuales cuelga de una percha teórica apropiada.

Los tres fenómenos son:

[A] La distribución aparentemente isotrópica en gran escala de las fuentes de radiación observadas desde la tierra.

[B] El desplazamiento sistemático hacia el rojo del espectro de las radiaciones recibidas de fuentes muy distantes. Advertido primero por Vesto Slipher antes de 1920, este fenómeno fue confirmado y establecido sólidamente por Edwin Hubble antes de 1930.

[C] La radiación térmica de cuerpo negro casi perfectamente isotrópica en la cual vivimos sumidos, cuyo origen no puede imputarse a ninguna fuente de radiación definida. Detectada primero en 1964 por Penzias y Wilson a frecuencias de microondas, ha sido medida luego para el espectro entero desde satélites en órbita fuera de la atmósfera, en alucinante concordancia con la ley de Planck.

El fenómeno **[A]** respalda el elemento **[1]** del modelo estándar, pero solo si damos por supuesto el principio copernicano conforme al cual la tierra está situada en un lugar cualquiera.[1] Solo si la distribución de las fuentes de radiación es isotrópica en torno a *cada punto* del universo se puede concluir, por el teorema de Schur, que esa distribución es homogénea.[2] La aplicación de este teorema de la geometría diferencial en el presente caso supone, por cierto, un vínculo entre distribución de las masas y geometría, esto es, una teoría geométrica de la gravitación por lo menos afín a la de Einstein. Adviértase que la ciencia humana, atada como está al sistema solar, no puede invocar respaldo empírico para el principio copernicano sin caer en un círculo vicioso. Este principio debe verse más bien como un prerrequisito a priori de la cosmología empírica, pues si ocupásemos una posición especial en el espacio no podríamos aprender mucho acerca del universo en su conjunto mediante los sentidos. Una cosmología científica depende de la posibilidad de inferir conclusiones universales de datos recogidos localmente.[3]

El principio copernicano implica que el fenómeno **[C]** también es universal: la radiación de fondo debe ser isotrópica en torno a cada punto del espacio. En tal caso, el elemento **[4]** del modelo estándar

está respaldado por el fenómeno **[C]** entendido a la luz de la fisica cuántica. La familia de teorías físicas que reunimos bajo esta denominación está bien corroborada por toda clase de experimentos y se justifica aplicarla a la radiación universal de fondo. Según la ley de Planck el espectro de la radiación observada corresponde a una radiación térmica sumamente fría (temperatura cercana a 3 K o –270°C). La física cuántica de partículas explica la omnipresencia de una radiación tal como el vestigio de una radiación térmica que estuvo acoplada a la materia y que, por lo tanto, igual que ésta, llenaba homogéneamente todo el espacio. Pero la radiación sólo puede estar acoplada a la materia si es muchísimo más caliente que 3 K. Por lo tanto, la explicación cuántica supone que la radiación de fondo se ha enfriado enormemente desde la época en que radiación y materia estaban acopladas. Tal enfriamiento es una consecuencia inevitable de la expansión del universo —segundo elemento del modelo estándar— la cual, por lo demás, explica también por qué, en cierto momento, la radiación se desacopló de la materia. Desde ese momento, la radiación desacoplada llena el universo como un fantasma que, aunque cada día está más frío, sigue siendo un testigo fiel de lo que ocurrió entonces. Si suponemos que el universo tuvo una época en que era extremadamente caliente y denso, la física cuántica de partículas explica también con bastante exactitud la abundancia relativa de hidrógeno y helio en el universo actual, la cual puede considerarse como un fenómeno **[D]** que también da respaldo al modelo estándar.

Pero el principal testigo del elemento **[2]**, vale decir, de la expansión del universo, es, por cierto, el fenómeno **[B]**, esto es, el corrimiento hacia el rojo del espectro de la luz recibida de fuentes remotas. Cuando los astrónomos estadounidenses, poco después de 1910, empezaron a medir el corrimiento de los espectros estelares hacia el azul y el rojo lo concibieron naturalmente como un efecto Doppler, indicativo de las velocidades radiales con que las estrellas se acercan a la tierra o se alejan de ella. Cuando se comprobó que las fuentes más lejanas regularmente exhiben un corrimiento al rojo que además, al parecer, es mayor cuanto más dista la fuente, se concluyó que esas fuentes debían clasificarse como *galaxias*,[4] esto es, como sistemas separados de estrellas —“universos islas”— análogos a nuestra Vía

Láctea, y que en su gran mayoría se están alejando de ésta y, en virtud del principio copernicano, también unas de otras, a velocidades proporcionales a sus distancias mutuas. Se han propuesto otras explicaciones para este fenómeno, pero con escaso éxito. Con todo, la visión de un universo que se expande frenéticamente se aviene tan mal con las ideas de estabilidad cósmica tradicionalmente cultivadas por la filosofía y la ciencia europeas que sorprende la facilidad y comparativa rapidez con que ella ha logrado una aceptación general. La circunstancia histórica puede haber ayudado: sin duda, el nombre y aun el concepto de *Big Bang* estaban bien adaptados al ambiente de la guerra fría, cuando todos creíamos que nuestro mundo humano podía acabar cualquier día en una catástrofe nuclear. Pero hay también un ingrediente de la historia interna de la física que, a mi modo de ver, fue decisivo: Las ecuaciones de campo de Einstein (ECE) que, desde su publicación en 1915 han venido obteniendo un apoyo cada vez más firme de observaciones practicadas en el sistema solar, admiten soluciones que prescriben la expansión del universo desde un estado de densidad altísima —en rigor, infinita— puramente por la fuerza de su geometría. Pertenecen a una familia de soluciones de las ECE descubierta por el matemático ruso Alexander Friedmann (1922, 1924) y redescubierta independientemente, bajo supuestos algo más amplios, por el astrónomo belga Georges Lemaître (1927). Su trabajo pasó bastante desapercibido hasta 1930, cuando Arthur Eddington, autoridad mundialmente reconocida en este campo, se percató de la existencia del trabajo de Lemaître, se refirió a él en un artículo suyo (1930, p. 668), lo hizo traducir al inglés y obtuvo su publicación en las *Monthly Notices of the Royal Astronomical Society*. Es en esta guisa que la teoría de la gravitación de Einstein o *relatividad general* constituye el espinazo teórico del modelo estándar, el tercer elemento que unifica a los otros tres. Si se postula una distribución global homogénea de la materia, conforme al elemento **[1],** las ECE admiten varias alternativas: o bien (i) la densidad global homogénea ha disminuido en un tiempo finito desde un estado inicial de densidad infinita y seguirá disminuyendo indefinidamente, o bien (ii) ella ha aumentado indefinidamente hasta alcanzar su valor actual y alcanzará un estado final de densidad infinita dentro de un tiempo finito, o bien

(iii) la densidad homogénea de la materia disminuye desde el infinito hasta un valor finito, alcanzado el cual, vuelve a crecer hasta el infinito, todo ello en un tiempo finito. Los elementos **[2]** y **[4]** del modelo estándar excluyen la alternativa (ii), pero son compatibles con la (i) y la (iii). Así, dado el elemento **[1]**, los elementos **[2]** y **[4]** se siguen del elemento **[3]**: dada la homogeneidad de la materia, la expansión del universo desde un estado caliente y denso es sólo una consecuencia de la geometría dinámica de la relatividad general.

En este capítulo trataré de justificar lo que he dicho sobre las soluciones de Friedmann repasando rápida pero ceñidamente el razonamiento contenido en sus artículos. Para hacerlo enteramente claro y persuasivo, necesitaría demostrar unos cuantos teoremas difíciles de geometría diferencial y transcribir muchos cómputos fáciles pero tediosos. No tengo aquí espacio para ello. Sin embargo, confío en que, al señalar explicitamente cada paso que omito desarrollar y desplegar a modo de ejemplo al menos algunos cómputos, mi exposición arrojará luz sobre la índole y el alcance de los modelos cosmológicos de Friedmann y servirá de guía para estudiar sus escritos. Pero antes de abordar su razonamiento debo dar al lector una idea de su significado físico. Con este propósito bosquejaré primero la motivación y los rasgos principales de la relatividad general (§ 5.2) y las tres soluciones exactas de las ecuaciones de campo de Einstein descubiertas antes que las de Friedmann (§§ 5.3–5.5).

Sigo las convenciones siguientes:

1ª Los índices que son letras de nuestro alfabeto recorren el conjunto {0,1,2,3}; los índices que son letras griegas recorren el conjunto {1,2,3}.[5]

2ª Conforme a la llamada *convención de Einstein*, se sobrentiende la suma sobre los índices repetidos dentro de un mismo término; por ejemplo, $u^\mu v_\mu + B^k{}_k = u^1 v_1 + u^2 v_2 + u^3 v_3 + B^0{}_0 + B^1{}_1 + B^2{}_2 + B^3{}_3$.

3ª Las unidades de medir se eligen de modo que la velocidad de la luz $c = 1$.

5.2 La invención de la relatividad general

Einstein sentó las bases de lo que hoy llamamos relatividad especial en un artículo presentado a la revista *Annalen der Physik* el 30 de junio de 1905 (1905*r*). El requisito de invariancia bajo las transformaciones de Lorentz que esta teoría impone a las ecuaciones de la física no es satisfecho por la ley de gravedad de Newton. En la víspera de Navidad de 1907 Einstein le escribió a Habicht que estaba trabajando en una nueva teoría de la gravedad. Ese mismo año Minkowski había presentado por primera vez su interpretación de la relatividad especial como física del espaciotiempo.[6] Einstein se demoró en apreciar la reformulación de su nueva física por su viejo maestro, pero cuando finalmente Sommerfeld le abrió los ojos ella le condujo a una concepción de la gravedad completamente inesperada e innovadora.[7]

Minkowski concibe la escena del acontecer físico como una ☆variedad cuatridimensional☆ $\mathcal{M}$, homeomorfa☆ a $\mathbb{R}^4$ y provista de una ☆métrica riemanniana☆ que llamaré $\boldsymbol{\eta}$.[8] La métrica $\boldsymbol{\eta}$ determina una partición del ☆espacio tangente☆ $\mathrm{T}_P\mathcal{M}$ en cada punto $P \in \mathcal{M}$ en tres clases de vectores. Un vector $v_P \in \mathrm{T}_P\mathcal{M}$ es *temporaloide* si $\boldsymbol{\eta}_P(v_P,v_P) > 0$; v_P es *espacialoide* si $\boldsymbol{\eta}_P(v_P,v_P) < 0$; v_P es *luminoide* o *nulo* si $\boldsymbol{\eta}_P(v_P,v_P) = 0$. Todo espacio tangente $\mathrm{T}_P\mathcal{M}$ satisface la siguiente *condición de ortogonalidad*: si $v_P \in \mathrm{T}_P\mathcal{M}$ es temporaloide y $w_P \in \mathrm{T}_P\mathcal{M}$ no es el vector 0, entonces $\boldsymbol{\eta}_P(v_P,w_P) = 0$ solo si w_P es espacialoide.[9] Las curvas☆ de $\mathcal{M}$ cuyas tangentes☆ pertenecen todas a una de estas clases reciben el nombre de esa clase. Estos nombres se adoptaron para expresar las siguientes connotaciones físicas: una curva espacialoide une eventos[10] que podrían ser simultáneos conforme a una definición físicamente viable del orden temporal (quiero decir, una definición con arreglo a la cual la muerte de una persona no pueda en ningún caso preceder a su nacimiento); una curva luminoide une eventos que podrían ser el punto de emisión y el de recepción de una señal luminosa transmitida en el vacío; una curva temporaloide puede ser la *cosmolínea* de un punto masa, pasando por los eventos sucesivos de su historia. A la variedad riemanniana $\langle\mathcal{M},\boldsymbol{\eta}\rangle$, que Minkowski llamó *el mundo* (*die Welt*), la llamamos hoy *espaciotiempo de*

Minkowski. En esta palestra las partículas libres trazan geodésicas☆ temporaloides[11] y los pulsos de luz transmitidos en el vacío trazan geodésicas luminoides. Esta doble condición garantiza la validez de los dos principios en que Einstein funda la relatividad especial en 1905, el *principio de relatividad* y el *principio de la constancia de la velocidad de la luz* (véase la § 1.2). Ello implica que el movimiento inercial y la propagación de la luz en el vacío dependen exclusivamente de la geometría del espaciotiempo. De este modo, en la relatividad especial, el papel de guía del movimiento no forzado que Newton confirió al espacio y el tiempo absoluto es asumido simplemente por el espaciotiempo absoluto de Minkowski, otro "fantasma que actúa sobre todas las cosas pero sobre el cual las cosas no reaccionan".[12] La comprobación de que esto es así no podía caerle bien a Einstein, que a la sazón adhería a las críticas que Mach (1883) le hizo a Newton y era partidario de referir todos los fenómenos físicos a la materia presente y actuante alrededor.

En el verano de 1912 Einstein tuvo la ocurrencia de que la geometría del espaciotiempo contenía la clave para la recta comprensión de los fenómenos gravitacionales. Ya su visión de 1907, que más tarde describió como "la idea más feliz de mi vida", lo había seguramente preparado para tenerla: un hombre que cae libremente desde el techo de un edificio muy alto disfruta de un breve período de ingravidez, junto con los objetos —llaves o monedas— que saque de sus bolsillos y suelte en torno suyo mientras va cayendo; ese hombre "tiene por tanto el derecho a interpretar su propio estado como 'en reposo'" (véase la cita en la p. 72). Originalmente, Einstein expresó esta visión mediante el concepto de equivalencia entre marcos de referencia: *No es posible establecer, mediante experimentos realizados en el interior de laboratorio cerrado* —digamos, una cápsula espacial—*, si este cae libremente en un campo gravitacional uniforme o si reposa relativamente a un sistema inercial*.[13] Con este *principio de equivalencia*, Einstein entiende haber resuelto el enigma de la igualdad proclamada por Newton entre la masa con que cada cuerpo resiste a la aceleración por una fuerza externa y la masa con que responde a las solicitaciones del campo gravitacional. La masa inercial y la masa gravitacional son cuantitativamente iguales porque son esencialmente

idénticas, por cuanto, en definitiva, la gravitación es inercia. Pero el principio de equivalencia por sí mismo no nos dice nada acerca de los campos gravitacionales reales, que —salvo a una distancia infinita de toda materia— no son nunca uniformes. Con todo, la "idea más feliz" de Einstein había asociado la caída libre al movimiento inercial, que Minkowski había vinculado a la geometría física. En el espaciotiempo cada partícula masiva tiene, en cada instante de su historia, una determinada velocidad espaciotemporal o *cosmovelocidad*, representada por la tangente a su cosmolínea en ese instante. Si la partícula no está sometida a la acción de fuerzas externas, dicha tangente mantiene un tamaño y dirección constantes a lo largo de la cosmolínea de la partícula. Podría decirse que la partícula, cuando ninguna fuerza la impele en una u otra dirección, se orienta por la geometría del espaciotiempo. Esta es la clave de la gravedad como la concibe Einstein. Había proclamado la equivalencia del movimiento inercial con la caída libre en un campo gravitacional uniforme. Simpatizaba activamente con la idea de Mach de que los fenómenos de la inercia, no menos que los de la gravedad, dependen de la distribución de la materia en el universo. Ahora vio una oportunidad de hacerla operante concibiendo al movimiento inercial como un caso límite de caída libre. Se decidió a tratar la geometría del espaciotiempo como la única fuente de orientación para las partículas en caída libre en un campo gravitacional cualquiera, a la vez que subordinaba la geometría a la distribución universal de la materia.

La relatividad general concibe la escena del acontecer físico como una variedad cuatridimensional no necesariamente homeomorfa a $\mathbb{R}^4$, que llamaré $\mathcal{U}$ (por 'universo'). $\mathcal{U}$ está provista de una métrica riemanniana $\mathbf{g}$. $\mathbf{g}$ satisface en cada espacio tangente a $\mathcal{U}$ la misma condición de ortogonalidad impuesta arriba a la métrica minkowskiana $\boldsymbol{\eta}$.[14] $\mathbf{g}$ está vinculada a la distribución de la materia y la energía no-gravitacional por las ecuaciones de campo de Einstein (en adelante: ECE). En su forma más general (Einstein 1917b), las ECE se pueden escribir así:

$$R_{ij} - \tfrac{1}{2} g_{ij} R + \lambda g_{ij} = -\kappa T_{ij} \tag{1}$$

En el lado izquierdo de (1), las g_{ij} son los componentes la métrica **g** y las R_{ij} son los componentes del tensor de Ricci☆, relativos a una carta☆ de $\mathcal{U}$; $R = R^k{}_k$ es el ☆escalar de curvatura☆ correspondiente a **g** y λ es un parámetro ajustable, llamado *la constante cosmológica*. En el lado derecho, κ es la constante gravitacional[15] y los T_{ij} son los componentes del tensor de energía, que representan la densidad, las tensiones y cizalles[16] localmente operantes y el impulso (o momento cinético) de la materia y la energía gravitacional presente en cada punto del dominio de la carta elegida. Poniendo $\lambda = 0$, obtenemos un sistema de ecuaciones equivalente a la versión original de las ECE.[17] La misma forma de las ecuaciones (1) implica que **g** solo excepcionalmente puede ser plana, y será generalmente una métrica de curvatura variable. Sin embargo, en virtud de la condición de ortogonalidad enunciada arriba (y explicada en la nota 14), **g** coincide hasta las cantidades de primer orden con la métrica plana minkowskiana $\boldsymbol{\eta}$ en un entorno —generalmente pequeño— de cada punto de $\mathcal{U}$. Ello concuerda con el éxito de la relatividad especial en los laboratorios y permite extender a $\mathcal{U}$ la clasificación de las curvas en temporaloides, espacialoides y nulas.

Las ecuaciones (1) no son lineales y por eso son notoriamente difíciles de resolver. Además, por la misma razón, no se puede dar por descontado que las partículas en caída libre trazarán en todos los casos cosmolíneas geodésicas, como deberían hacerlo si, conforme al programa de Einstein, fuesen guiadas por la métrica **g**. No se puede postular una "ley geodésica del movimiento" que simplemente prescriba que esto tiene que ser así, porque no toda distribución de masas que se muevan de cualquier modo arbitrario es compatible con el sistema no lineal (1). Por lo tanto, la ley del movimiento tiene que *derivarse* de las ECE. Se ha logrado hacerlo en varios casos especiales, pero una solución general del problema del movimiento de la relatividad general no se conoce y podría ser imposible (cf. la nota 39 al capítulo 1, p. 210). Sin embargo, esta dificultad no tiene por qué inquietarnos aquí, pues la sencillísima distribución de la materia presupuesta por las soluciones de Friedmann es uno de los casos en que sabemos que la ley geodésica se cumple exactamente.

Pero antes de abordar las soluciones de Friedmann de las ECE debo referirme a tres soluciones exactas que fueron descubiertas antes: la solución exterior de Schwarzschild (1916*a*) y las soluciones cosmológicas de Einstein (1917*b*) y de Sitter (1917*a*,*b*,*c*).

5.3 La solución de Schwarzschild para el espacio vacío

Karl Schwarzschild (1916*a*) halló la primera solución exacta de las ECE en su versión original —ecuaciones (2) en la nota 17 (p. 227)— muy poco después de su publicación.[18] Esta solución está especialmente adaptada a las necesidades de la astronomía del sistema solar. Los planetas se conciben como partículas de prueba —esto es, como objetos de masa tan insignificante que ella no afecta al campo gravitacional— que caen libremente en el espacio vacío alrededor de un cuerpo central considerado como la única fuente del campo. Schwarzschild tomó de Einstein (1915*h*) la hipótesis de que las cosmolíneas de tales partículas de prueba son geodésicas temporaloides y por lo tanto satisfacen el principio variacional

$$\delta\int ds = 0 \tag{5}$$

donde el ☆elemento de línea☆ *ds* está dado, en términos de las coordenadas —supuestamente globales[19]— x^0, x^1, x^2, x^3, por

$$ds^2 = g_{ij}dx^i dx^j \qquad (i,j = 0,\ 1,\ 2,\ 3) \tag{6}$$

Enseguida, procedió a resolver las ecuaciones (2) para los componentes g_{ij} de la métrica, bajo condiciones especiales adaptadas al sistema solar idealizado arriba descrito. Schwarzschild supone que la fuente gravitacional (i) no varía con el tiempo y (ii) está concentrada en el origen de las coordenadas espaciales x^1, x^2, x^3.[20] En el resto del es-

pacio los componentes T_{ij} del tensor de energía son todos constantes e iguales a 0, de modo que las ecuaciones (2) se reducen a:

$$R_{ij} = 0 \tag{7}$$

Schwarzschild supone además que (iii) una partícula de prueba infinitamente distante de la fuente se mueve uniformemente en línea recta, conforme al clásico principio de inercia. Los supuestos (i), (ii) y (iii) se traducen respectivamente en las siguientes condiciones que Schwarzschild (1916*a*, p. 190) repite, palabra por palabra, de Einstein (1915*h*, p. 833), pero yo doy en traducción libre al castellano:

S1 Los componentes de la métrica no dependen del tiempo x^0.
S2 Las ecuaciones $g_{0\rho} = g_{\rho 0} = 0$ valen exactamente para $\rho = 1, 2, 3$. (En el artículo de Einstein esta condición tiene el número 3.)
S3 La solución es espacialmente simétrica en torno al eje $x^1 = x^2 = x^3 = 0$; dicho de otro modo: se obtiene la misma solución si las coordenadas x^1, x^2, x^3 se someten a una transformación ortogonal o rotación. (En el artículo de Einstein esta condición tiene el número 2.)
S4 En el infinito espacial, esto es, a infinita distancia del eje de simetría, las g_{ik} convergen a 0 si $i \neq j$; de no ser así convergen a estos límites: $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$.

La condición S4 dice en efecto que el espaciotiempo es minkowskiano —y por ende plano— en el infinito espacial. La condición S2 significa que las ☆curvas paramétricas☆ de la coordenada temporal x^0 son ortogonales en todo lugar a las hipersuperficies x^0 = const. Esta condición se introdujo probablemente solo para facilitar los cálculos, pues carece de toda significación física en el vacío. En una solución genérica de las ECE esta condición de ortogonalidad no es necesariamente satisfecha por cualquier carta. Sin embargo, el solo hecho de que Schwarzschild produjese una solución exacta que satisface las cuatro condiciones demuestra que la condición S2 no es incompatible con las otras tres.

No necesitamos entrar en más detalles de la solución de Schwarzschild. Debo mencionar, sí, un importante teorema matemático demostrado pocos años más tarde por G. D. Birkhoff (1923). Las con-

diciones S1 y S4 se deducen ambas de la condición S3. Un campo gravitacional gobernado por las ECE e invariante bajo rotaciones en torno a un eje temporaloide es necesariamente estático y converge a la métrica plana de Minkowski en el infinito espacial.

*5.4 La solución cosmológica de Einstein (1917b).**

Las "Consideraciones cosmológicas" presentadas por Einstein a la academia de Berlín el 8 de febrero de 1917 marcan el nacimiento de la cosmología moderna. No responden a hechos recién descubiertos sino a la insatisfacción filosófica de Einstein con la solución de Schwarzschild. La exigencia de que el espaciotiempo sea plano en el infinito espacial (S4) implica que la métrica espaciotemporal impone una cosmolínea bien definida a una partícula de prueba aunque esta diste infinitamente de todas las fuentes de gravitación. Einstein creía que esto viola el "principio de Mach" (*Mach'sches Prinzip*), según el cual "el campo [métrico] está determinado *exhaustivamente* por las masas de los cuerpos" (1918*e*, p. 241). Asegurar este principio era a la sazón quizás el principal incentivo que movía a Einstein a enfocar la gravedad desde el punto de vista geometrodinámico.[21]

Einstein repudia en estos términos el comportamiento del campo de Schwarzschild en el infinito espacial: "En una teoría de la relatividad que sea consistente consigo misma no puede haber inercia *con respecto al 'espacio'*, sino solo inercia de las masas *respecto las unas de las otras*. Por lo tanto, si transporto una masa a una distancia suficientemente grande de todas las demás, su inercia debe caer a cero" (1917*b*, p. 145). Si el campo gravitacional es isotrópico en el

* Esta sección repite lo dicho al respecto en el capítulo 4. Quien haya leído ese capítulo hace poco puede omitirla. Conviene que lea, eso sí, las condiciones E1–E3, que he redactado en un vocabulario similar al que usaré para enunciar las condiciones postuladas por Friedmann y algo diferente del utilizado en el capitulo 4.

espacio en torno a cada punto, hay un sistema de coordenadas $x = (x^0,x^1,x^2,x^3)$, relativamente al cual la métrica espaciotemporal está dada por el ☆elemento de línea☆:

$$ds^2 = B(dx^0)^2 - A((dx^1)^2 + (dx^2)^2 + (dx^3)^2) \qquad (8)$$

Einstein muestra que, en tal caso, la masa inercial de una partícula cae a cero en el infinito espacial si y solo si A también se anula allí, mientras que B crece por sobre todo límite asignable. Según Einstein, "tales condiciones de frontera son absolutamente inadmisibles para el sistema de las estrellas fijas" (1917*b*, p. 146). Basándonos en el teorema de Birkhoff podemos ahora ir más lejos: las condiciones $A \to 0$ y $B \to \infty$ no solo están en conflicto con propiedades observables de las estrellas fijas, sino que son *lógicamente incompatibles* con la simetría postulada.

Se le ocurrió entonces a Einstein que una métrica riemanniana puede estar realizada en una variedad espacialmente compacta☆. En un espaciotiempo con esta propiedad no hay un infinito espacial y por tanto no hay que postular condiciones de frontera que se cumplan allí. Einstein se propuso hallar una solución de las ECE que cumpliese las siguientes condiciones:

E1 El espaciotiempo admite una foliación[22] canónica en hipersuperficies espacialoides compactas☆.
E2 El campo métrico es homogéneo e isotrópico sobre cada hipersuperficie de la foliación canónica.[23]
E3 Las hipersuperficies de la foliación canónica son isométricas entre sí.

La condición E1 permite prescindir de la postulación de condiciones de frontera en el infinito espacial. Supongamos que la materia y la radiación de hecho están distribuidas isotrópicamente en torno nuestro. Combinando este hecho con el principio "copernicano" según el cual cualquier lugar del espacio es aproximadamente igual al que nosotros ocupamos, se deduce la condición E2. La condición E3 dice en efecto que la métrica del espaciotiempo es estacionaria en el tiempo. Con ello, Einstein pretende dar una expresión idealizada a

los datos astronómicos que parecían indicarle que las estrellas se mueven al azar en todas direcciones a velocidades relativas muy inferiores a la velocidad de la luz. Tales datos provenían, por cierto, solo de observaciones de estrellas de nuestra galaxia. Curiosamente, si no rebasamos sus límites, no cabe sostener que la materia está distribuida isotrópicamente en torno a nosotros, pues casi todos sus astros están situados en una zona más bien angosta del cielo. Por tanto, sospecho que la condición E2 no refleja información astronómica sino más bien el deseo de reducir el problema planteado por las ECE a términos matemáticamente manejables.

Para satisfacer las tres condiciones a la vez Einstein reemplazó las ECE originales —ecuaciones (2) en la nota 17— con las ecuaciones (3), equivalentes a nuestras ecuaciones (1). Sin embargo, señala expresamente que "una curvatura positiva del espacio" —acorde con las condiciones E1 y E2— puede derivarse también de las ecuaciones (2) y que el término nuevo λg_{ij} (con $\lambda \neq 0$) insertado en las ecuaciones (3) se requiere únicamente para asegurar "una distribución cuasi-estática de la materia que refleje el hecho de las bajas velocidades estelares", o sea, la condición E3 (1917*b*, p. 152). Me cuesta imaginar que Einstein haya publicado esta provocativa observación sin haber antes calculado una solución no estática de las ecuaciones (2) con un espaciotiempo foliado en hipersuperficies espacialoides compactas que se pueda, por tanto, describir como un espacio universal finito pero ilimitado. Sea de ello lo que fuere, pienso como Barbour (1990, p. 63 n. 7) que el citado señalamiento de Einstein estimuló la investigación de Friedmann.

5.5 El universo vacío de Willem de Sitter

Para completar este panorama de los antecedentes de la obra de Friedmann, debo mencionar asimismo la solución de las ecuaciones (1) comunicada por Willem de Sitter a la academia de Amsterdam el 31 de marzo de 1917. De Sitter supone que el espaciotiempo es una variedad compacta homeomorfa a $\mathbf{S}^4$, la hipersuperficie esférica de

cuatro dimensiones (incrustable en $\mathbb{R}^5$).[24] De Sitter explica que esta "idea de hacer que el mundo cuatridimensional sea esférico para evitar la necesidad de fijar condiciones de frontera" le había sido sugerida varios meses antes por Paul Ehrenfest, en una conversación privada, pero no fue desarrollada en ese momento (de Sitter 1917*a*, p. 1219 n.1). Esto me hace pensar que de Sitter solo apreció las potencialidades de la sugerencia de Ehrenfest una vez que Einstein hizo público su propio mundo cilíndrico, homeomorfo a $\mathbb{R} \times \mathbf{S}^3$. De Sitter critica a Einstein por atribuirle al tiempo una condición privilegiada.

> Pues hablar de *el* mundo tridimensional, si no equivale a introducir un tiempo absoluto, al menos implica la hipótesis de que en cada punto del espacio cuatridimensional hay una coordenada determinada x_4 que es preferible a todas las otras para representar el "tiempo", y que siempre y en todo punto esta sola coordenada se elige efectivamente como "tiempo". Una diferencia tan fundamental entre el tiempo y las coordenadas espaciales parece contradecir un tanto la completa simetría de las ecuaciones de campo y de las ecuaciones del movimiento (ecuaciones de la línea geodésica) con respecto a las cuatro variables.
>
> (de Sitter 1917*a*, p. 1223)

La solución de Willem de Sitter supone un universo vacío. Einstein, a quien de Sitter comunicó lo esencial de su artículo antes de publicarlo, le había escrito el 24 de marzo: "En mi opinión sería insatisfactorio que se pudiera pensar en un mundo sin materia. El campo g^{ij} debe más bien *estar condicionado por la materia y no poder subsistir sin ella*. Este es el núcleo de lo que entiendo por el postulado de la relatividad de la inercia."[25] Por su parte, de Sitter subraya que "toda la materia que conocemos, estrellas, nebulosas, cúmulos, etc." no basta para asegurar ese postulado; "los seguidores de Mach se ven pues compelidos a suponer la existencia de aún más materia" (1917*c*, p. 5). En el universo cilíndrico de Einstein, la masa total de esta materia hipotética "es tan enormemente grande que, comparada con ella, toda la materia conocida por nosotros es totalmente insignificante" (de Sitter 1917*a*, p. 1219). De Sitter prescinde de toda esta materia adicional y trata "toda la materia conocida por nosotros" como partículas de prueba. La geometría de su solución implica que "la frecuencia de las

vibraciones luminosas disminuye con el incremento de la distancia desde el origen de las coordenadas". Por lo tanto "las líneas en los espectros de las estrellas y nebulosas muy distantes deben desplazarse sistemáticamente hacia el rojo, generando la apariencia espuria de una velocidad radial positiva (*giving rise to a spurious positive radial velocity*)" (1917*c*, p. 26). Velocidades radiales de –311, +925 y +1185 km/s habían sido recientemente registradas para tres galaxias por más de un observador. El promedio era +600 km/s. Comenta de Sitter: "Por cierto, este resultado derivado de solo tres nebulosas prácticamente no tiene valor. Sin embargo, si una observación continuada confirmase el hecho de que las nebulosas espirales exhiben sistemáticamente velocidades radiales positivas, esto sería ciertamente una indicación de que la hipótesis B [esto es, el modelo esférico propuesto por de Sitter—R.T.] ha de preferirse a la hipótesis A [el modelo cilíndrico de Einstein]" (1917*c*, p. 28).[26]

5.6 Las soluciones de Friedmann

El aporte de Alexander Friedmann a la cosmología está contenido en dos artículos, "Sobre la curvatura del espacio" (1922) y "Sobre la posibilidad de un mundo con curvatura negativa constante del espacio" (1924), escritos en alemán y publicados en la *Zeitschrift für Physik*. En el primero de ellos, Friedmann sostiene que, suponiendo que el tensor de energía tome la forma más simple, el universo cilíndrico de Einstein y el esférico de de Sitter son las únicas soluciones de las ecuaciones (1) que cumplen las condiciones E1, E2 y E3;[27] pero que, si se descarta la condición E3, hay un continuo de soluciones que satisfacen las condiciones E1 y E2. En el segundo artículo, eliminó el requisito de compacidad incluido en la condición E1 y produjo otra familia no numerable de soluciones.

Para entender la obra de Friedmann hay que prestar atención a la índole peculiar de las ECE. Estas ecuaciones expresan una relación bien definida entre tres campos tensoriales simétricos de rango 2, a saber, el tensor de energía, la métrica del espaciotiempo y el tensor

de Ricci determinado por la métrica, y un campo escalar también determinado por la métrica, a saber, el escalar de curvatura R. En las ecuaciones (1) los campos tensoriales están representados por su ☆componentes covariantes☆ relativos a una carta espaciotemporal cualquiera x.[28] Por las ecuaciones (7.25) y (7.39), sabemos que tanto el escalar de curvatura R como los componentes R_{ij} del tensor de Ricci envuelven las derivadas segundas de los componentes métricos g_{ij} respecto a las coordenadas x^k. Por tanto, si los componentes T_{ij} del tensor de energía están dados, tenemos diez ecuaciones diferenciales de segundo orden, que bastan para determinar las diez incógnitas g_{ij}.[29] Por cierto, es prácticamente imposible hallar una expresión moderadamente realista del tensor de energía, por ejemplo, sobre el volumen de una colmena de abejas durante todo un día de verano, pero cuando se trata del universo entero la tarea se facilita mucho idealizando audazmente y tomando promedios. Así, Friedmann supone que una nube de polvo incoherente en caída libre llena el universo. Si el campo vectorial V, con componentes covariantes V_i, asigna a cada punto del espaciotiempo la cosmovelocidad de la nube material en ese punto,[30] los componentes del tensor de energía del polvo están dados por la ecuación

$$T_{ij} = \rho V_i V_j \tag{9}$$

El campo escalar ρ asigna a cada punto espaciotemporal la densidad de la nube en ese punto.[31] Como aun esta condición resulta demasiado amplia, Friedmann, siguiendo el ejemplo de Einstein (1917*b*) impone una restricción más al movimiento de la materia: “La materia está […] en reposo relativo; o, dicho menos rigurosamente, las velocidades relativas de la materia son insignificantes comparadas con la velocidad de la luz” (Friedmann 1922, p. 378). Entonces, según Friedmann, las ecuaciones (9) se pueden escribir así:

$$\begin{gathered} T_{00} = \rho g_{00} \\ T_{ij} = 0 \text{ si } i \neq 0 \text{ o } j \neq 0 \end{gathered} \tag{10}$$

Las ecuaciones (10) valen si, además de postular (*a*) que las velocidades de las partículas de polvo en términos de las coordenadas ele-

gidas son prácticamente nulas, suponemos (*b*) que los componentes métricos $g_{0j} = 0$ para $j \neq 0$. Friedmann introduce este supuesto suplementario algunos párrafos más adelante (p. 379; equivale a la condición G3 que enunciaré enseguida). Bajo estas condiciones, las ecuaciones (10) se deducen fácilmente de las ecuaciones (9). En virtud del postulado (*a*), los componentes espacialoides de la cosmovelocidad, V^1, V^2 y V^3, son por doquier iguales a 0. Por lo tanto, en virtud del supuesto (*b*), $V_j = g_{ij}V^i = 0$ si $j \neq 0$. Por lo tanto, $T_{ij} = \rho V_i V_j = 0$ si $i \neq 0$ o $j \neq 0$. Por la definición de V, $V^iV^jg_{ij} = 1$ (nota 30). Como $V^j = g_{0j} = 0$ cuando quiera que $j \neq 0$, tenemos que $V^0V^0g_{00} = 1$. Ahora bien, $V_0 = g_{0j}V^j = g_{00}V^0$. Por lo tanto, $T_{00} = \rho V_0 V_0 = \rho g_{00}V^0V^0g_{00} = \rho g_{00}$.

Las ecuaciones (1), (9) y (10) son un epítome de las condiciones físicas del problema que Friedmann se propuso resolver. Pero para hallar soluciones de las ecuaciones (1) tuvo que imponer más condiciones "concernientes al carácter general, por así decir geométrico del mundo" (p. 378). Estas condiciones geométricas se pueden parafrasear así:[32]

G1 El espaciotiempo admite una coordenada temporal global t, que lo folia en hipersuperficies t = const.

G2 La restricción de la métrica espaciotemporal **g** a una dada hipersuperficie $\mathfrak{H}_T$ en que la coordenada t tiene por doquier el mismo valor T induce en $\mathfrak{H}_T$ una métrica riemanniana propiamente tal (vale decir, ☆positiva definida☆) de curvatura constante, la cual puede, sí, depender del tiempo.

G3. Cada hipersuperficie de la foliación mencionada en la condición G1 es ortogonal en cada punto a la ☆curva paramétrica☆ de la coordenada t que pasa por ese punto.

La condición G1 recoge la idea tradicional de que todos los eventos del universo pueden clasificarse de algún modo en clases mutuamente excluyentes de eventos simultáneos. Al parecer, Friedmann consideraba esto tan obvio que no lo menciona expresamente, pero está implícito por cierto en su foliación del universo entero en hipersuperficies espacialoides (conforme a la condición G2). Hawking (1969) demostró que las soluciones de las ECE satisfacen la condición G1 siempre que sean *causalmente estables*, es decir, siempre que el

respectivo espaciotiempo no admita curvas temporaloides cerradas ni las admitiría en el caso de una pequeña perturbación de la métrica.

La condición G2 equivale a la aseveración aparentemente más débil de que cada clase de eventos simultáneos llena un espacio homogéneo —no contiene puntos privilegiados— e isotrópico —no hay direcciones preferidas en torno a ningún punto. Como ya señalé, esta aseveración se deduce de la isotropía observada de la materia y la radiación, combinada con el principio copernicano, el teorema de Schur y la supuesta dependencia de la geometría espaciotemporal respecto a la distribución de la energía (véase la nota 2, p. 225). En relación con esto, conviene tener presente que la isotropía observada no se percibe en un dado instante *T* sobre la hipersuperficie espacialoide lisa $\mathfrak{H}_T$ que circunda al observador, sino sobre el cono de luz pretérito del observador en *T*, esto es, sobre la hipersuperficie nula generada por los rayos de luz que el observador recibe o podría recibir en ese instante. Ahora bien, el teorema de Schur no es aplicable a esta superficie, que no es isotrópica en torno a cualquier punto, sino solo en torno al punto en que corta la cosmolínea del observador, y además no es una variedad tridimensional lisa, pues no es lisa en ese punto.[33]

Respecto a la condición G3, Friedmann dice que no puede ofrecer ninguna justificación física o filosófica para adoptarla y que "ella sirve exclusivamente para simplificar los cómputos" (p. 379). Como no toda solución de las ECE cumple con la condición G3, esta débil justificación parecería restarle significación física a la obra de Friedmann. Él mismo quizás lo sentía así, pues se apresura a añadir, en el mismo párrafo numerado en que enuncia la condición G3, que los modelos cosmológicos de Einstein y de Sitter ambos satisfacen los requisitos geométricos que ha postulado. Pero H. P. Robertson demostró más tarde a partir de un teorema clásico de la geometría diferencial que las condiciones G1 y G2 implican la condición G3. De hecho, Robertson invoca en lugar de G2 la siguiente premisa aparentemente más débil:

G2′ Cada hipersuperficie espacialoide $\mathfrak{H}_t$ en que la coordenada temporal global *t* sea constante es isotrópica en torno a cada uno de sus puntos.

Por el teorema de Schur, G2′ implica que, para cada valor admisible de t, $\mathfrak{H}_t$ es una variedad tridimensional homogénea. Como $\mathfrak{H}_t$ es espacialoide, cualquier vector $v \neq 0$ tangente a $\mathfrak{H}_t$ satisface la desigualdad $\mathbf{g}(v,v) < 0$, donde $\mathbf{g}$ es la métrica del espaciotiempo. Por lo tanto, la restricción $\mathbf{g}\restriction\mathfrak{H}_t$ de $\mathbf{g}$ a $\mathfrak{H}_t$ determina una métrica riemanniana propiamente tal $\mathfrak{g}_t$ sobre $\mathfrak{H}_t$, definida por $\mathfrak{g}_t = -(\mathbf{g}\restriction\mathfrak{H}_t)$. Conforme a un teorema enunciado por el mismo Riemann, una variedad riemanniana propiamente tal homogénea tiene necesariamente curvatura constante positiva, negativa o nula (Riemann 1854, pp. 144s.). Hasta aquí no hemos ido más allá de las condiciones G1 y G2. El paso decisivo a la condición G3 se basa en un teorema demostrado por Fubini (1904), que enunciaré sin probarlo.

Tras postular G1, Robertson formula G2′ como sigue:

> Requerimos que para cualquier observador estacionario ("cuerpo de prueba") en este universo idealizado todas las direcciones (espaciales) en torno suyo sean totalmente equivalentes, en el sentido de que sea incapaz de distinguirlas mediante una propiedad intrínseca del espaciotiempo; [requerimos] además que sea igualmente incapaz de detectar una diferencia entre sus observaciones y las de cualquier otro observador contemporáneo.
>
> (Robertson 1929, p. 823)[34]

Por lo tanto, el espaciotiempo debe admitir la acción de un grupo de movimientos[35] que aplique sobre sí misma toda hipersuperficie que sea el lugar geométrico de eventos contemporáneos, ya sea haciéndola rotar alrededor de un punto cualquiera, ya sea trasladando un punto selecto cualquiera a cualquier otro. Como las rotaciones y traslaciones arbitrarias en una variedad tridimensional suponen seis grados de libertad, el grupo en cuestión tiene que ser un grupo de seis parámetros. El teorema de Fubini dice que un grupo de movimientos de seis parámetros que actúa sobre una variedad riemanniana cuatridimensional actúa transitivamente como grupo de movimientos sobre cualquier miembro de una familia de hipersuperficies con las siguientes propiedades:

F1 Todas las hipersuperficies son ortogonales a una congruencia☆ de geodésicas, que llamaré $\mathscr{C}$.

F2 Cada par de hipersuperficies intercepta segmentos de igual longitud en cada geodésica de $\mathcal{C}$.

F3. Si $\mathfrak{H}$ y $\mathfrak{H}'$ son dos hipersuperficies de la familia y la aplicación biyectiva $f: \mathfrak{H} \to \mathfrak{H}'$ envía cada punto $p \in \mathfrak{H}$ al punto donde $\mathfrak{H}'$ corta la geodésica de $\mathcal{C}$ que pasa por p, entonces f es una aplicación conforme.[36]

La propiedad F1 de Fubini implica la condición G3 de Friedmann si, pero solo si las curvas paramétricas de la coordenada temporal global t son geodésicas. Esto se sigue necesariamente de las hipótesis físicas de Friedmann si las curvas paramétricas de t coinciden con las cosmolíneas de la materia, o pueden hacerse coincidir con ellas mediante una reparametrización; pues dichas cosmolíneas son geodésicas temporaloides, por las ecuaciones (1) y (9). Para asegurar que así sea, basta elegir las coordenadas x^μ ($\mu = 1, 2, 3$) de modo que su valor permanezca constante a lo largo de la cosmolínea de cada partícula del polvo cósmico, esto es, de modo que las coordenadas espaciales constituyan un sistema de referencia comóvil con la materia. Me parece que eso es precisamente lo que Friedmann quiere decir cuando postula, antes de derivar las ecuaciones (10), que "la materia…está en reposo relativo" (1922, p. 378).[37] Entonces, en virtud de F2, la coordenada temporal t tiene que ser proporcional al tiempo propio. Referido a tales coordenadas, el elemento de línea toma la forma

$$ds^2 = M^2 dt^2 - S(t)^2 \mathfrak{g}_{\mu\nu} dx^\mu dx^\nu \qquad (11)$$

donde (i) M es el factor de proporcionalidad entre la coordenada t y el tiempo propio a lo largo de cada cosmolínea de la materia; (ii) los $\mathfrak{g}_{\mu\nu}$ son los componentes relativos a las coordenadas espaciales x^μ de la métrica riemanniana propiamente tal $\mathbf{\mathfrak{g}}_{+} = -(\boldsymbol{g} \restriction \mathfrak{H}_{+})$ inducida por $\mathbf{g}$ sobre una hipersuperficie espacialoide $\mathfrak{H}_{+}$ elegida arbitrariamente como estándar, en la cual t toma un valor constante t_{+}, y (iii) $S(t)$ es el factor escalar de la aplicación conforme $f: \mathfrak{H}_{+} \to \mathfrak{H}_t$, para cada valor admisible de t (véase la nota 36, p. 233). Para simplificar la notación, en lo sucesivo escribo $\mathbf{\mathfrak{g}}$ en vez de $\mathbf{\mathfrak{g}}_{+}$. Obsérvese que $S(t_1) = 1$, por definición, y que en general, debido a la homogeneidad e isotropía de

las hipersuperficies $\mathfrak{H}_t$, $S(t)$ depende de t únicamente. Como $S(t) \neq 0$ para cada valor admisible de t, $S(t)$ no puede cambiar de signo. Supongamos que es siempre positivo. Debido a la isotropía de la hipersuperficie $\mathfrak{H}_+$ en torno a cada punto, las coordenadas espaciales se pueden elegir de modo que, escribiendo r por x^1, θ por x^2 y φ por x^3, los componentes de $\mathfrak{g}$ sean:

$$\mathfrak{g}_{11} = \frac{1}{1-kr^2} \qquad \mathfrak{g}_{22} = r^2 \qquad \mathfrak{g}_{33} = r^2 \operatorname{sen}^2\theta \tag{12}$$

$$\mathfrak{g}_{\mu\nu} = 0 \text{ if } \mu \neq \nu$$

donde k es una constante dependiente de la curvatura constante de $\mathfrak{H}_+$, que ponemos igual a 1, –1 o 0, según que esa curvatura sea positiva, negativa o nula, respectivamente. (En la terminología hoy habitual, el modelo cosmológico se dice 'cerrado' en el primer caso, 'abierto' en el segundo y 'plano' en el tercero.) Elegimos la coordenada t de tal modo que $M = 1$. Referida a la carta con coordenadas (t,r,θ,φ), el elemento de línea (6) toma la forma siguiente:

$$ds^2 = dt^2 - S(t)^2\left(\frac{dr^2}{1-kr^2} + r^2 d\theta^2 + r^2 \operatorname{sen}^2\theta d\varphi^2\right) \tag{13}$$

Este es el llamado *elemento de línea de Robertson-Walker*, en honor de H. P. Robertson y A. G. Walker, que lo descubrieron, cada uno por separado, hacia 1935. Aunque difiere de los elementos de línea que figuran expresamente en los dos artículos de Friedmann, se lo usa normalmente en la literatura actual y seguiré esta práctica.[38] Los componentes covariantes g_{ij} de la métrica aparecen desplegados en la matriz siguiente

$$\left(g_{ij}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\dfrac{S(t)^2}{1-kr^2} & 0 & 0 \\ 0 & 0 & -S(t)^2 r^2 & 0 \\ 0 & 0 & 0 & -S(t)^2 r^2 \operatorname{sen}^2\theta \end{pmatrix} \tag{14}$$

y los componentes contravariantes g^{ij} son, claro está, los elementos de la matriz inversa:

$$\left(g^{ij}\right) = \left(g_{ij}\right)^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\dfrac{1-kr^2}{S(t)^2} & 0 & 0 \\ 0 & 0 & -\dfrac{1}{S(t)^2 r^2} & 0 \\ 0 & 0 & 0 & -\dfrac{1}{S(t)^2 r^2 \operatorname{sen}^2\theta} \end{pmatrix} \qquad (15)$$

La presencia del factor $S(t)^2$, dependiente del tiempo, en cada uno de los $g_{\mu\mu}$ indica que en un espaciotiempo con elemento de línea de Robertson-Walker la separación espacialoide entre eventos simultáneos a lo largo de dos cosmolíneas dadas puede aumentar o disminuir con el transcurso del tiempo *debido a la geometría del universo*, tal como la distancia entre lugares de igual latitud sobre dos meridianos terrestres decrece continuamente del ecuador a los polos *debido a la geometría del globo terráqueo.* El cambio de frecuencia de la luz enviada desde una de esas cosmolíneas a la otra puede deducirse directamente de la ecuación (13) si suponemos que las señales luminosas describen geodésicas luminoides (nulas) y que su frecuencia no es afectada por accidentes físicos mientras viajan. Supongamos que una partícula material A emite luz en el instante t_1 durante un breve lapso de tiempo Δt_1, y que esta luz es recibida por la partícula B en el instante t_2 durante el intervalo Δt_2. Como el número de ondas luminosas recibidas es igual al número de ondas emitidas, la frecuencia de emisión ν_1 y la frecuencia de recepción ν_2 satisfacen la ecuación $\nu_1 \Delta t_1 = \nu_2 \Delta t_2$. Sobre la geodésica luminoide trazada por una señal luminosa

$$ds^2 = dt^2 - S(t)^2 d\sigma^2 = 0 \qquad (16)$$

donde $d\sigma$ es el elemento de línea sobre la hipersuperficie estándar $\mathfrak{H}_{+}$ ($d\sigma^2 = g_{\mu\nu}dx^{\mu}dx^{\nu}$). Por lo tanto, la distancia espacial σ recorrida por la señal entre A y B está dada por

$$\sigma = \int_{t_1}^{t_2} \frac{dt}{S(t)} \tag{17}$$

Como σ no depende de t, también está dada por

$$\sigma = \int_{t_1+\Delta t}^{t_2+\Delta t} \frac{dt}{S(t)} \tag{18}$$

donde Δt es pequeño. Substrayendo (17) de (18) y descartando las cantidades pequeñas $|S(t_1+\Delta t) - S(t_1)|$ y $|S(t_2+\Delta t) - S(t_2)|$, obtenemos

$$\frac{\nu_2}{\nu_1} = \frac{\Delta t_1}{\Delta t_2} = \frac{S(t_1)}{S(t_2)} \tag{19}$$

El cambio de frecuencia z se define convencionalmente como la razón entre la pérdida de frecuencia y la frecuencia observada, y en consecuencia está dado por

$$z = \frac{\nu_1 - \nu_2}{\nu_2} = \frac{S(t_2)}{S(t_1)} - 1 \tag{20}$$

Si $S(t_2) > S(t_1)$ —esto es, si el universo se ha "expandido"[39] en el tiempo trascurrido entre t_1 y t_2—, $z > 0$ y constituye por lo tanto un "corrimiento al rojo".

Para derivar el elemento de línea (13) y el cambio de frecuencia (20) hemos invocado tres hipótesis básicas de la teoría de la gravedad de Einstein, a saber, (i) que la luz describe geodésicas nulas (usamos esto en la ecuación (16)), (ii) que el polvo cósmico describe geodésicas temporaloides (esto nos permitió igualar la condición G3 con F1), y (iii) que la geometría del espaciotiempo depende de la distribución de las fuentes de gravedad (de modo que G2′ se infiera, por el principio copernicano y el teorema de Schur, de la isotropía de la materia y la

energía observada aquí y ahora). Hay que destacar, por otra parte, que hasta este punto *las* ECE *no han contribuido nada a nuestro razonamiento.* Para ver qué luz arrojan sobre la "dinámica del universo en expansión" —según lo expresa el elemento **[3]** del modelo estándar conforme a la caracterización de Peebles (p. 103)— hay que resolver las ecuaciones (1) para el caso especial en que los componentes g_{ij} de la métrica están dados por la matriz (14) y los componentes T_{ij} del tensor de energía por las ecuaciones (10). Con este fin hay que explicitar completamente los componentes del tensor de Ricci en términos de los componentes de la métrica. Un vistazo a las ecuaciones (7.40) sugiere que nos aguarda un trabajo espantoso. Sin embargo, en el caso extraordinariamente simple de una métrica de Friedmann, la tarea no es en absoluto difícil, aunque envuelve cómputos tediosos. Los presentaré en un par de casos, a modo de ejemplo. Exhorto al lector a que tome papel y lápiz y calcule el resto. Escribo S, S' y S'', en vez de $S(t)$, dS/dt y d^2S/dt^2, respectivamente.

Tenemos que $g_{00} = g^{00} = 1$ y que $g_{ik} = g^{ik} = 0$ cuando $i \neq k$, de modo que $\partial g_{0k}/\partial x^j = 0$ para todos los índices k y j, y que $\partial g_{\mu\nu}/\partial x^j = 0$ para todo j si $\mu \neq \nu$. Esto implica que, en el caso considerado, la mayoría de los Γ^i_{kh} que figuran en las ecuaciones (7.40) —y que las ecuaciones (7.39) expresan en términos de los g_{ij}— son iguales a 0. Por ejemplo,

$$\Gamma^0_{00} = \tfrac{1}{2} g^{0h}\left(\frac{\partial g_{h0}}{\partial x^0} + \frac{\partial g_{0h}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^h}\right) = \tfrac{1}{2} g^{00}\left(\frac{\partial g_{h0}}{\partial x^0}\right) = 0 \qquad (21)$$

De hecho, las únicas excepciones son estas:

$$
\begin{aligned}
\Gamma^1_{01} &= \tfrac{1}{2} g^{11}\left(\frac{\partial g_{10}}{\partial x^1} + \frac{\partial g_{11}}{\partial x^0} - \frac{\partial g_{01}}{\partial x^1}\right) \\
&= \tfrac{1}{2} g^{11}\left(\frac{\partial g_{11}}{\partial x^0}\right) = -\frac{\tfrac{1}{2} S^{-2}}{\mathfrak{g}_{11}}(-2SS'\mathfrak{g}_{11}) = \frac{S'}{S} \\
\Gamma^2_{02} &= \tfrac{1}{2} g^{22}\left(\frac{\partial g_{22}}{\partial x^0}\right) = -\frac{\tfrac{1}{2} S^{-2}}{\mathfrak{g}_{22}}(-2SS'\mathfrak{g}_{22}) = \frac{S'}{S} \\
\Gamma^3_{03} &= \tfrac{1}{2} g^{33}\left(\frac{\partial g_{33}}{\partial x^0}\right) = -\frac{\tfrac{1}{2} S^{-2}}{\mathfrak{g}_{33}}(-2SS'\mathfrak{g}_{33}) = \frac{S'}{S} \\
\Gamma^0_{11} &= \tfrac{1}{2} g^{00}\left(-\frac{\partial g_{11}}{\partial x^0}\right) = -SS'\mathfrak{g}_{11} \\
\Gamma^0_{22} &= \tfrac{1}{2} g^{00}\left(-\frac{\partial g_{22}}{\partial x^0}\right) = -SS'\mathfrak{g}_{22} \\
\Gamma^0_{33} &= \tfrac{1}{2} g^{00}\left(-\frac{\partial g_{33}}{\partial x^0}\right) = -SS'\mathfrak{g}_{33}
\end{aligned}
\tag{22}
$$

y los componentes distintos de 0 de la ☆conexión de Levi-Civita☆ de la métrica $\mathfrak{g}$, pues, como el lector comprobará, si los componentes de esta conexión se designan con $\mathfrak{G}^\lambda_{\mu\nu}$, donde $\lambda, \mu, \nu \in \{1,2,3\}$,

$$
\Gamma^\lambda_{\mu\nu} = \mathfrak{G}^\lambda_{\mu\nu} \tag{23}
$$

Gracias a que casi todos los $\Gamma^\lambda_{\mu\nu}$ se anulan, se simplifica enormemente el cómputo de los componentes del tensor de Ricci. Por ejemplo,

$$
\begin{aligned}
R_{00} &= R^{k}_{ijk} = \frac{\partial \Gamma^{k}_{0k}}{\partial x^{0}} - \frac{\partial \Gamma^{k}_{00}}{\partial x^{k}} + \Gamma^{s}_{0k}\Gamma^{k}_{s0} - \Gamma^{s}_{00}\Gamma^{k}_{sk} = \\
&= 3\frac{d}{dt}\left(\frac{S'}{S}\right) + \left(\Gamma^{\mu}_{0\mu}\right)^{2} = 3\left(\frac{S''S - (S')^{2}}{S^{2}} + \left(\frac{S'}{S}\right)^{2}\right) \qquad (24)\\
&= \frac{3S''}{S}
\end{aligned}
$$

En forma análoga calculamos que

$$R_{0\mu} = 0 \qquad (25)$$

y

$$R_{\mu\nu} = \Re_{\mu\nu} - \mathfrak{g}_{\mu\nu}(SS'' + 2(S')^2) \qquad (26)$$

donde los $\Re_{\mu\nu}$ son los componentes del tensor de Ricci construido a partir de la métrica $\mathfrak{g}$. Se puede demostrar que en una variedad riemanniana propiamente tal $\langle \mathcal{M},\mathfrak{g}\rangle$ con n dimensiones y curvatura constante k, $\Re_{\mu\nu} = (1 - n)k\mathfrak{g}_{\mu\nu}$ (Weinberg 1972, pp. 382s.). Por lo tanto, las ecuaciones (26) pueden reescribirse así

$$R_{\mu\nu} = -\,\mathfrak{g}_{\mu\nu}(SS'' + 2(S')^2 + 2k) \qquad (26')$$

Usando las ecuaciones (15), (25) y (26′) calculamos fácilmente el escalar de curvatura:

$$
\begin{aligned}
R &= g^{ij}R_{ij} = R_{00} + g^{11}R_{11} + g^{22}R_{22} + g^{33}R_{33} \\
&= \frac{3S''}{S} - \left(SS'' + 2(S')^{2} + 2k\right)\left(-\frac{3}{S^{2}}\right) \\
&= 6\left(\frac{S''}{S} + \frac{(S')^{2} + k}{S^{2}}\right)
\end{aligned}
\qquad (27)
$$

Ahora tenemos con qué reemplazar R_{ij}, R, g_{ij} y T_{ij} en las ecuaciones (1). Si omitimos las ecuaciones en que ambos lados resultan ser idénticamente 0, las ECE se reducen a:

$$\frac{3}{S^2}\left((S')^2 + k\right) - \lambda = \kappa\rho \tag{28}$$

$$\frac{1}{S^2}\left(2SS'' + (S')^2 + k\right) - \lambda = 0 \tag{29}$$

Estas son las ecuaciones del campo gravitacional de los universos de Friedmann. El significado físico de la densidad ρ y la constante cosmológica λ se pone de manifiesto si combinamos ambas ecuaciones para eliminar el término $\left((S')^2 + k\right)$. Multiplicando ambos lados de la ecuación (29) por 3, substrayendo de la ecuación (28) el resultado y rearreglando los términos, obtenemos:

$$\frac{S''}{S} = \left(\frac{\lambda}{3} - \frac{\kappa\rho}{6}\right) \tag{30}$$

Por lo tanto, con λ positiva se apresura el crecimiento del factor escalar *S*, como si actuase una fuerza newtoniana repulsiva entre las partículas materiales; mientras que ρ positiva tiene el efecto contrario y aparece entonces como una fuente de atracción universal.

Debido a la homogeneidad e isotropía de las hipersuperficies $\mathfrak{H}_t$, la densidad ρ varía únicamente con *t*. Conforme a los supuestos acerca del tensor de energía adoptados por Friedmann, la conservación de la energía implica que

$$\rho S^3 = \text{const.} = \rho_0(S_0)^3 \qquad (31)^{40}$$

donde, como es habitual, distingo con el subíndice 0 los valores actuales de cantidades dependientes del tiempo. Usando la ecuación (31) para hacer reemplazos en la ecuación (28) obtenemos:

$$(S')^2 = \frac{\kappa\rho_0\left(S_0\right)^3}{3s} + \frac{\lambda S^2}{3} - k \tag{32}$$

La ecuación (30) puede reordenarse así:

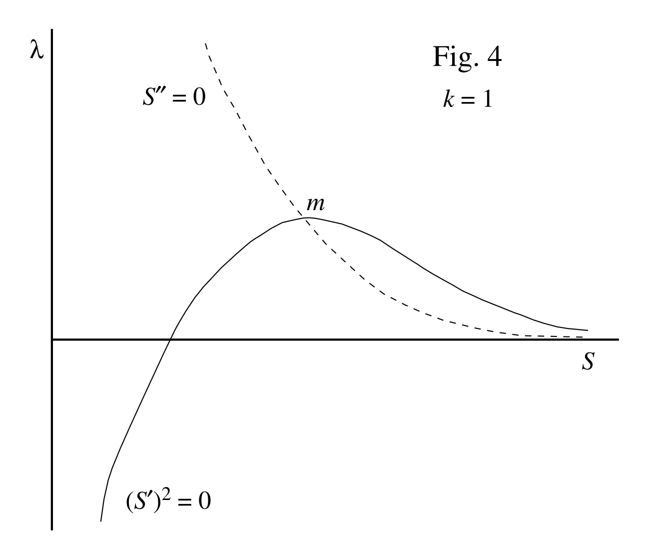
$$S'' = -S\left(\frac{\kappa\rho}{6} - \frac{\lambda}{3}\right) \tag{33}$$

Como la densidad ρ es positiva siempre, la ecuación (33) implica que, si la constante cosmológica λ es negativa o igual a 0, o si κρ es siempre mayor que 2λ, S'' siempre es negativa. En tal caso, el factor escalar S no puede tener un mínimo y converge a 0 cuando t se aproxima a cierto valor finito t_N (donde el subíndice N puede leerse *Nada* o *Nihil*). En vista de la ecuación (32), es claro que

$$\begin{aligned} t_N &= \int_0^N dt = \int_{S(0)}^0 \frac{dt}{dS}dS = \int_{S(0)}^0 \frac{dS}{S'} = \\ &= \int_{S(0)}^0 dS\sqrt{\frac{3S}{\lambda S^3 - 3kS + \kappa\rho_0(S_0)^3}} \end{aligned} \tag{34}$$

Si $t = 0$ ahora, esto es, si $S(0) = S_0 > 0$, y el valor actual de S' es positivo, t_N es un número finito negativo al que la coordenada t se acerca en la dirección del pasado. La ecuación (31) implica que, cuando S se aproxima a 0, la densidad ρ crece por sobre todo límite asignable.[41] De las ecuaciones (27), (28) y (30) deducimos que el escalar de curvatura $R = \kappa\rho + 4\lambda$. Así, también R crece más allá de todo límite cuando t se aproxima a t_N. Por lo tanto, en un universo de Friedmann como el que estamos considerando, la coordenada temporal t_N no corresponde a ningún evento. El número real t_N obtenido evaluando la integral elíptica en el lado derecho de la ecuación (34) no puede hallarse en el camino de la coordenada temporal global t, pero si $S' > 0$ ahora, t_N es, por cierto, el ínfimo de esa coordenada (su cota inferior máxima). Decir que $|t_N|$ es "el tiempo trascurrido desde la creación del mundo" ("die Zeit seit der Erschaffung der Welt"— Friedmann 1922, p. 384) suena bonito y no hay inconveniente en hablar de este modo, siempre que no se pierda de vista que en un universo de Friedmann —y en cualquier modelo cosmológico de la relatividad general— *todo evento ha sido precedido por otros* y, por

lo tanto, *cada evento E tiene un pasado* (aunque, claro está, si $t(E)$ es apenas mayor que t_N, ese pasado será brevísimo).



Robertson y Noonan (1968, pp. 376s.) distinguen varios tipos de universos de Friedmann, a saber, el universo oscilante, dos clases de universos monotónicos y varios casos ligados a un valor especial de λ. Sus rasgos característicos pueden leerse en las Figs. 4 y 5,[42] que presentan la curva $(S')^2 = 0$ en el plano λ-S en los dos casos $k = 1$ y $k \in \{0,-1\}$. En ambos casos, la curva $(S')^2 = 0$ se acerca asintóticamente al eje positivo de los valores de S y al eje negativo de los valores de λ. Si $k = 1$, la curva $(S')^2 = 0$ llega a un punto $m = \langle S_m,\lambda_m\rangle$ tal que $\lambda_m > 0$, a partir del cual λ decrece monotónicamente sobre la curva a medida que crece S. Si $k \leq 0$, el valor de λ sobre $(S')^2 = 0$ aumenta monótonicamente con S pero nunca pasa de 0 (no cruza el eje de las S). Como el cuadrado de S' no puede ser negativo, ningún punto situado bajo la curva $(S')^2 = 0$ representa un modelo cosmológico posible. Como λ es constante dentro de cada modelo, el punto que representa a un modelo dado solo puede desplazarse en una recta horizontal, hacia la izquierda si $S' < 0$, y hacia la derecha si $S' > 0$. Cuando $S' = 0$, el incremento o disminución del factor escalar S se detiene, al menos momentáneamente. S es constante solo

si también $S'' = 0$. La condición $S' = S'' = 0$ es a la vez necesaria y suficiente para que un universo de Friedmann sea estático, pues, como se infiere de la ecuación (30), todas las derivadas de S respecto del tiempo se anulan si se cumple esta doble condición.[44] Las figuras muestran que la curva $S'' = 0$ corta la curva $(S')^2 = 0$ solo si $k = 1$ y $\lambda = \lambda_m$. El modelo cosmológico de Einstein (1917*b*) es el único universo de Friedmann que llena estos requisitos.

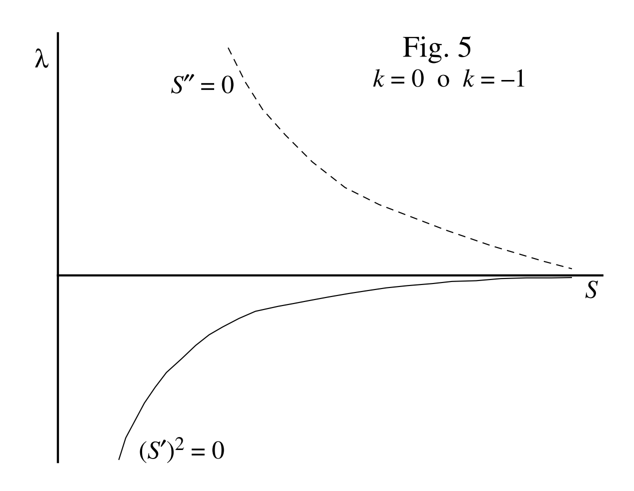


Fig. 5

Cualquier otro universo de Friedmann se expande o contrae continuamente. El *universo oscilante* está confinado entre el eje de las λ y la curva $(S')^2 = 0$. Se expande desde $S = 0$ hasta que S alcanza un máximo y luego se recontrae hasta que $S = 0$ nuevamente. El universo oscilante tiene que satisfacer una de las cuatro condiciones siguientes:

(i) $k = -1$ y $\lambda < 0$
(ii) $k = 0$ y $\lambda < 0$
(iii) $k = 1$ y $\lambda < 0$
(iv) $k = 1$, $0 < \lambda < \lambda_m$ y $S < S_m$

El *universo monotónico de la primera clase* o bien se expande monotónicamente desde $S = 0$, sin límite, para siempre, o bien se

contrae monotónicamente desde el infinito hasta $S = 0$. El incremento o disminución de S no puede revertirse porque $S' \neq 0$ siempre. Esta clase de universo tiene que satisfacer una de las tres condiciones siguientes:

(i) $k = -1$ y $\lambda \geq 0$
(ii) $k = 0$ y $\lambda \geq 0$
(iii) $k = 1$ y $\lambda > \lambda_m$

El *universo monotónico de la segunda clase* puede contraerse monotónicamente desde cualquier valor dado de $S > S_m$ hasta un valor mínimo de $S > S_m$, pero de ahí en adelante se expande monotónicamente, sin límite, para siempre. Esta clase de universo tiene que satisfacer la condición siguiente:

$$k = 1,\ 0 < \lambda < \lambda_m \text{ y } S > S_m$$

Ninguno de los tipos mencionados incluye modelos en que $\lambda = \lambda_m$, el valor especial arriba descrito. Como dije, esto solo puede ocurrir si $k = 1$, esto es, si las hipersuperficies espacialoides $\mathfrak{H}_t$ son compactas y tienen curvatura constante positiva. Hay cinco casos posibles:

(i) Si $S < S_m$ y $S' > 0$, el modelo se expande desde $S = 0$ y se aproxima asintóticamente al punto m.
(ii) Si $S < S_m$ y $S' < 0$, el modelo se contrae hasta $S = 0$ y estuvo asintóticamente próximo a m en el pasado.
(iii) Si $S > S_m$ y $S' < 0$, el modelo se contrae monotónicamente y se aproxima asintóticamente a m.
(iv) Si $S < S_m$ y $S' > 0$, el modelo estuvo asintóticamente próximo a m en el pasado y se expande monotónicamente para siempre.
(v) Si $S = S_m$, el modelo permanece en m; por lo tanto, $S' = S'' = 0$, y el modelo es estático.

El caso (v) corresponde a la solución de las ecuaciones (1) propuesta en las "Consideraciones cosmológicas" de Einstein (1917*b*). Su lugar geométrico en el plano λ-S consta de un solo punto m, en cada entorno del cual hay puntos que corresponden a universos en expansión o en contracción. Por lo tanto, esta solución es inestable: la más

mínima perturbación convierte el modelo estático en uno dinámico. Comprobamos así que, contra lo que Einstein esperaba, la inserción en las ECE del término λg_{ij}, con $\lambda \neq 0$, está lejos de asegurar la inmutabilidad del universo en el contexto de la relatividad general.

Hay acuerdo general en que la constante cosmológica λ es, o bien igual 0, o bien muy pequeña, mucho más pequeña que λ_m, el máximo valor de λ que es compatible con $(S')^2 = 0$. En vista del fenómeno **[B]**, el universo en que vivimos estaría expandiéndose. Por lo tanto, si se lo puede representar mediante un modelo de Friedmann, este tiene que ser o bien un universo oscilante que actualmente se halle en su fase expansiva, o bien un universo expansivo monotónico de la primera clase, con $k = 0$ o $k = -1$. Las ecuaciones (28) y (29) relacionan el valor de k con el valor actual ρ_0 de la densidad media del universo y otras dos cantidades observables, a saber, la "constante" de Hubble H_0, definida como el valor actual del cociente (S'/S), y el parámetro de deceleración q_0, esto es, el valor actual de la cantidad $-(SS''/(S')^2)$, y también con la constante cosmológica λ, observable en principio aunque al parecer es demasiado pequeña para que afecte significativamente a nuestros instrumentos de observación.[44] Por la ecuación (28)

$$\rho_0 = \frac{3}{\kappa}\left(\left(\frac{S_0'}{S_0}\right)^2 + \frac{k}{S_0^{\,2}}\right) - \frac{\lambda}{\kappa} = \\ = \frac{3}{\kappa}\left(H_0^{\,2} + \frac{k}{S_0^{\,2}} - \frac{\lambda}{3}\right) \tag{35}$$

Claramente hay un valor crítico ρ_c, tal que $k = S_0^2(\rho_0 - \rho_c)$, a saber:

$$\rho_c = \frac{3}{\kappa}\left(H_0^{\,2} - \frac{\lambda}{3}\right) \tag{36}$$

$k = 1$ si y solo si $\rho_0 > \rho_c$. De la ecuación (29) inferimos que

$$\begin{aligned}\frac{k}{S_0^{\ 2}} &= \lambda - \left(\frac{S_0'}{S_0}\right)^2 - \frac{2S_0S_0''}{S_0^{\ 2}} \\ &= \lambda - \left(\frac{S_0'}{S_0}\right)^2 - \frac{2S_0S_0''}{S_0^{\ 2}}\left(\frac{S_0}{S_0'}\right)^2\left(\frac{S_0'}{S_0}\right)^2 \\ &= \lambda - (1-2q_0)H_0^{\ 2}\end{aligned} \tag{37}$$

Si λ es insignificante o igual a 0, las ecuaciones (35), (36) y (37) implican que:

$$\frac{\rho_0}{\rho_c} = \frac{H_0^{\ 2} - (1-2q_0)H_0^{\ 2}}{H_0^{\ 2}} = 2q_0 \tag{38}$$

Por lo tanto, si $q_0 > ½$, las hipersuperficies espacialoides $\mathfrak{H}_t$ tienen curvatura positiva y el universo de Friedmann es espacialmente compacto como el universo cilíndrico de Einstein (1917*b*,) pero si $q_0 \leq ½$ las hipersuperficies $\mathfrak{H}_t$ son planas o tienen curvatura negativa, y el universo de Friedmann es espacialmente infinito durante toda su historia, desde su mismo inicio en el llamado *Big Bang*.

No podemos considerar aquí los métodos empleados para medir H_0 y q_0. Por eso, en vez de examinar más de cerca el significado empírico de estas cantidades, me referiré, para terminar, a los horizontes. Aunque Friedmann no alude a ellos, son un rasgo esencial de sus modelos cosmológicos. Distinguimos entre *horizontes de partículas* y *horizontes de eventos.* Siguiendo a Rindler (1956), defino ambos conceptos para un espaciotiempo relativista $\langle \mathcal{U}, \mathbf{g} \rangle$ que admite una coordenada temporal universal t y en que las cosmolíneas de la materia forman una congruencia $\mathcal{C}$. $\mathcal{C}$ puede tratarse como variedad tridimensional de un modo obvio.[45] El *dominio causal* de un punto $p \in \mathcal{U}$, es el conjunto de todos los puntos $x \in \mathcal{U}$ unidos a p por una curva luminoide o temporaloide; el *pasado causal* de p comprende los puntos x del dominio causal de p tales que $t(x) < t(p)$; el *futuro causal* de p comprende los puntos x del dominio causal de p tales que $t(p) <$

$t(x)$. Digo que una cosmolínea $\gamma \in \mathcal{C}$ está más allá del horizonte de un punto $p \in \mathcal{U}$, si el pasado causal de p con contiene ningún punto del camino de γ. El *horizonte de partículas* de p es la frontera (en la variedad $\mathcal{C}$) del conjunto de cosmolíneas de $\mathcal{C}$ que están más allá del horizonte de p. Evidentemente, una partícula material cuya cosmolínea está más allá del horizonte de partículas p no puede ser la fuente de una señal luminosa o de otra forma de transferencia de energía que ejerza influencia sobre un evento en p. Digo que un punto $p \in \mathcal{U}$ está más allá del horizonte de una cosmolínea $\gamma \in \mathcal{C}$ si el futuro causal de p no contiene ningún punto del camino de γ. El *horizonte de eventos* de γ es la frontera (en la variedad $\mathcal{U}$) del conjunto de puntos de $\mathcal{U}$ que están más allá del horizonte de γ. Evidentemente, un evento cuya ubicación en el espaciotiempo esté más allá del horizonte de eventos de γ no puede ser la fuente de una señal luminosa o de otra forma de transferencia de energía que ejerza influencia sobre la historia de una partícula material cuya cosmolínea sea γ. En los modelos de la relatividad general es corriente que haya horizontes no vacíos de partículas y de eventos. Los agujeros negros está envueltos por horizontes de eventos. Hay horizontes de partículas en todo universo de Friedmann que se expande desde un límite temporal t_N tal que $\lim_{t \to t_N} S(t) = 0$. Digamos que el espaciotiempo $\langle \mathcal{U}, \mathbf{g} \rangle$ sea un universo de Friedmann de este tipo. Sea $\gamma \in \mathcal{C}$ la cosmolínea geodésica de una partícula de polvo cósmico. El horizonte de partículas de cualquier evento p en el camino de γ divide a $\mathcal{C}$ en dos componentes no vacíos, uno de los cuales, que llamaré $\mathcal{C}(p,\gamma)$, contiene a γ, mientras que el otro contiene todas las cosmolíneas de la materia que están más allá del horizonte de p. Dejemos que p recorra el camino de γ. Resulta entonces que, a medida que $t(p)$ se aproxima a t_N, $\mathcal{C}(p,\gamma)$ se contrae a γ. Por lo tanto, en el universo de Friedmann descrito, cada mínima porción de la materia surge en el *Big Bang* en completo aislamiento de todas las demás y de ahí en adelante adquiere de momento en momento un círculo creciente de conocidos, esto es, de partículas con las que ya ha interactuado. En la medida en que nuestro universo se parece, a grandes rasgos, a este modelo, nuestro actual horizonte de partículas está siendo penetrado ahora mismo, desde direcciones opuestas, por nuevas galaxias que nunca antes tuvie-

ron la oportunidad de interactuar con nosotros y cuya distancia mutua es demasiado grande para que hayan interactuado entre sí. Algunos físicos se sorprenden del gran parecido de estas galaxias, cuya radiación demuestra que tienen prácticamente la misma composición química y un perfil que las sitúa —de acuerdo con sus respectivas circunstancias— en el esquema de la evolución galáctica y estelar. Pero ¿qué hay de notable en esto? Solo implica que, en contra de la opinión de los filósofos, la causalidad no es el cemento del universo, al menos no de un universo que se asemeje en alguna medida a un modelo cosmológico relativista de cierto tipo. Pero las cosas no tienen que frotarse entre sí para formar parte de un mismo mundo. Más bien, tienen que ser suficientemente parecidas de entrada para que puedan reaccionar a su presencia mutua (y frotarse, llegado el caso). En la relatividad general, como en toda gran teoría de la física matemática, la estructura monta la escena y define las condiciones para la causalidad eficiente. En un universo en que hay horizontes, la primacía de la estructura luce con más esplendor.

6

La geometría decimonónica

En el siglo XIX, la geometría, como casi todas las disciplinas académicas, tuvo un período de crecimiento con proporciones de cataclismo: su contenido y su interna diversidad aumentan al punto que hacerla casi irreconocible; su método axiomático, tan celebrado desde la antigüedad, alcanza finalmente genuina suficiencia lógica, y se sientan las bases para que la familiar geometría de Euclides sea reemplazada, en la descripción de los fenómenos físicos, con el sistema maravillosamente flexible de Riemann. Los filósofos modernos de todas las tendencias —Descartes y Hobbes, Spinoza y Locke, Hume y Kant— veían en la geometría euclídea el paradigma de la certeza. Cuando se redujo de súbito a no ser más que una subespecie dentro de una vasta familia de teorías matemáticas se derrumbaron algunas ilusiones —por ejemplo, respecto a la posibilidad de una ciencia no menos segura de la moral[1]—, precipitando importantes cambios en la concepción filosófica del conocimiento humano. Repasaré aquí aquellos aspectos de la geometría decimonónica que más interesan desde este punto de vista, aludiendo de paso a su significación filosófica.

6.1 La geometría lobachevskiana

Euclides (siglos IV y III a.C.) puso a la cabeza de sus *Elementos* una serie de "definiciones" (ὅροι) —vgr. "un punto es aquello que no tiene partes"— y "nociones comunes" (κοιναὶ ἔννοιαι) —vgr. "Si se suman iguales a iguales, las sumas son iguales"—, así como cinco "peticiones" (αἰτήματα). Supuestamente, estos enunciados transmiten toda la información necesaria para inferir los teoremas y resolver los problemas de la geometría, pero de hecho no es así. De todos modos, para que las pruebas de Euclides sean concluyentes, hay que

aceptar las peticiones, comúnmente llamadas 'postulados' en castellano. Algunas de ellas son estrictamente prácticas: "1. Trazar una línea recta de cualquier punto a cualquier punto. [...] 3. Trazar un círculo con cualquier centro y cualquier radio." Sin embargo, la quinta suena más bien como la aseveración de un hecho. No obstante, se la puede leer sin dificultad como receta para construir triángulos. Dado un segmento *PQ*, hemos de trazar una recta *a* por *P* y una recta *b* por *Q*, de modo que *a* y *b* estén situadas en el mismo plano; se nos pide aceptar que, si los ángulos que *a* y *b* forman a uno de los lados de *PQ* suman menos de dos rectos, las rectas *a* y *b* se cortan en un punto *R* a ese mismo lado de *PQ*, formando el triángulo *PQR*. Esta petición se llama el *postulado de Euclides* (Fig. 6).

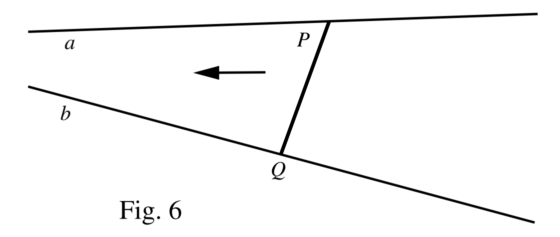


Fig. 6

Durante la progresiva eclipse de la inteligencia que acompañó a las conquistas de Roma, se perdió el sentido de la libertad matemática que tenía Euclides y los filósofos y matemáticos se acostumbraron a creer que la geometría tenía que basarse en fundamentos evidentes. Ahora bien, si *a* es perpendicular y *b* es casi perpendicular a *PQ*, *a* y *b* se acercan la una a la otra muy lentamente a un lado de *PQ* y está lejos de ser evidente que a la larga tengan que cortarse en algún punto a ese lado. Desde luego, la hipérbole se acerca indefinidamente a sus asíntotas y sin embargo puede demostrarse que no las corta jamás. En el trascurso de los siglos, varios autores exigieron —e intentaron— una demostración del postulado de Euclides. John Wallis (1616–1703) la derivó del supuesto de que hay polígonos de diverso tamaño que tienen la misma forma. Pero este supuesto tiene que demostrarse a su vez. Girolamo Saccheri (1667–1733) ensayó una reducción al absurdo; dedujo de la negación del postulado de Euclides una larga serie de proposiciones, hasta llegar a una que según él

“repugna a la naturaleza” de la línea recta.[2] Pero la idea que Saccheri se hacía de esta “naturaleza” provenía de la geometría euclídea, y su argumento contenía una petición de principio.

Hacia 1820, Nikolai I. Lobachevsky (1793–1856) y Janos Bolyai (1802–1860) abordaron independientemente esta cuestión de un modo radicalmente nuevo. Lobachevsky basó en la negación del postulado de Euclides una geometría alternativa que llamó “imaginaria” e intentó verificar su validez a la escala astronómica.[3] Bolyai eliminó ese postulado del sistema de Euclides; el remanente constituye la “geometría absoluta”, que puede especificarse más añadiéndole ya sea el postulado de Euclides o su negación. Desde fines del siglo XVIII, Carl Friedrich Gauss (1777–1855) había estado trabajando en este mismo sentido, pero no publicó sus resultados por temor “a la gritería de los beocios”.[4] Como el primero en publicar fue Lobachevsky, es apropiado llamar *geometría lobachevskiana* al sistema de geometría resultante de la unión de la “geometría absoluta” y la negación del postulado de Euclides.

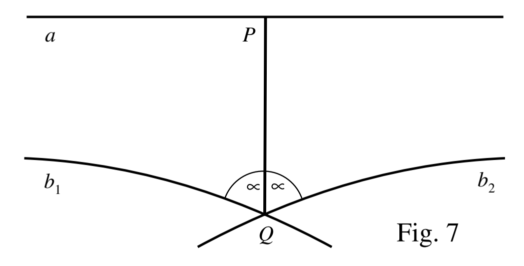


Fig. 7

La construcción usada arriba para explicar el postulado de Euclides también puede servirnos para elucidar su negación. Trácese por el punto *P* la recta *a* perpendicular al segmento *PQ* (Fig. 7). Si el postulado de Euclides es falso, hay incontables rectas coplanares[5] con *a* que pasan por *Q*, y forman ángulos agudos con *PQ* pero nunca cortan a la recta *a*. Considérese el conjunto de los números reales que expresan el tamaño —digamos, en grados— de esos ángulos agudos. Sea μ el ínfimo de este conjunto.[6] Evidentemente, $\mu > 0$. Hay exactamente dos rectas por *Q*, coplanares con *a*, que forman con *PQ* un ángulo de μ grados. Llamémolas b_1 y b_2. Ninguna de estas dos rectas

corta la recta *a*, pero *a* corta todas las líneas por *Q* que son coplanares con *a* y forman con *PQ* un ángulo menor que μ. Gauss, Lobachevsky and Bolyai —cada uno por su cuenta y sin que mediase ninguna comunicación entre ellos— llamaron a b_1 y b_2, *las paralelas* a la recta *a* por el punto *Q*. μ se llama el ángulo de paralelismo para el segmento *PQ*. Su tamaño depende de la longitud de *PQ*, y disminuye según esta aumenta. Supongamos que el ángulo de paralelismo para *PQ* sea igual a la mitad de un ángulo recto. En ese caso, b_1 y b_2 forman un ángulo recto con su vértice en *Q* y tenemos entonces dos rectas perpendiculares entre sí que yacen sobre el mismo plano que la recta *a* pero no la cortan en ningún punto.

La geometría de Lobachevsky abunda en teoremas sorprendentes (muchos de los cuales habían sido deducidos por Saccheri). Cito algunos: Los tres ángulos internos de un triángulo suman *menos* de dos rectos. La diferencia o *defecto* es proporcional al área del triángulo. Por lo tanto, en la geometría lobachevskiana los triángulos semejantes son siempre congruentes. Además, si el triángulo se divide en triángulos más pequeños, el defecto del todo es igual a la suma de los defectos de las partes. Como el defecto no puede ser mayor que dos rectos, hay un área máxima que ningún triángulo puede exceder. Si un cuadrilátero, por construcción, tiene tres ángulos rectos, el cuarto ángulo es necesariamente agudo. Por lo tanto, en la geometría lobachevskiana no hay rectángulos.

Hay una simple correspondencia formal entre las ecuaciones de la trigonometría lobachevskiana y las de la trigonometría esférica estándar. Basándose en ella, Lobachevsky sostuvo que cualquier contradicción que surja en su geometría tiene que corresponder inevitablemente a una contradicción en la geometría euclídea. Este sería el primer intento de probar la consistencia relativa de una teoría, esto es, de probar que ella es consistente a menos que otra teoría —cuya consistencia normalmente se da por sentada— sea inconsistente.

La geometría lobachevskiana recibió poca atención antes de 1865. La reacción de los filósofos, cuando finalmente se produjo, no fue unánime. Algunos vieron esta nueva geometría como un ejercicio formal de inferencia lógica, carente de significación física o filosófica, que utilizaba palabras familiares —como 'recta' y 'plano'— en

una acepción diferente de la habitual. Otros la celebraron como prueba suficiente de que, contrariando la influyente doctrina de Kant, la geometría euclídea no expresa prerrequisitos de la experiencia humana y la estructura geométrica del espacio físico es materia de investigación experimental. Otros, por último, reconocieron que una geometría no euclídea es una alternativa matemática legítima, pero señalaron que el diseño y la interpretación de los experimentos físicos generalmente presupone una geometría bien determinada y que el sistema de Euclides ya está instalado en esa función.

No importa cual fuese la opinión de los filósofos, para los matemáticos la geometría lobachevskiana probablemente no habría pasado de ser una curiosidad si no se hubieran encontrado sendos nichos para ella dentro de la geometría proyectiva y de la geometría diferencial, las dos vertientes principales de la geometría decimonónica (§6.2 y §6.5).

6.2 La geometría proyectiva

La geometría proyectiva no juega hoy un gran papel en las matemáticas, pero en el último tercio del siglo XIX llegó a ser sinónima de geometría moderna. Desargues (1591–1661) y Pascal (1623–1662) usaron métodos proyectivos, que fueron pronto desplazados por la geometría analítica de Descartes (1596–1650). Prosperaron nuevamente, sin embargo, después que Jean-Victor Poncelet (1788–1867) demostró que las propiedades proyectivas de las figuras suministraban medios de prueba no menos poderosos y ciertamente más geométricos que el procedimiento cartesiano de establecer y resolver ecuaciones entre números que representan puntos.

Las propiedades proyectivas se caracterizan porque son preservadas por las proyecciones. Tomemos, por ejemplo, dos planos Γ y Θ y un punto P fuera de ambos. Sea $\mathfrak{F}$ cualquier figura en el plano Γ. Tracemos rectas desde P por todos los puntos de $\mathfrak{F}$. La figura $\mathfrak{F}'$ formada por los puntos donde esas rectas cortan el plano Θ son la *proyección* de $\mathfrak{F}$ desde P sobre Θ. En general, $\mathfrak{F}'$ tiene otro tamaño

y otra forma que $\mathfrak{F}$ ($\mathfrak{F}$ y $\mathfrak{F}'$ no son congruentes, ni siquiera semejantes). Pero la proyección desde *P* sobre Θ de cualquier número de rectas que yacen en Γ y se cortan en ciertos puntos consiste generalmente en un número igual de rectas que yacen en Θ y se cortan respectivamente en las proyecciones de esos puntos. ¿Qué ocurre, sin embargo, si la recta que une a *P* con algún punto *Q* de Γ nunca corta a Θ, porque *PQ* está en un plano paralelo a Θ? Para obviar tales excepciones, la geometría proyectiva agregó a cada recta del espacio un *punto ideal* que comparte con todas sus paralelas. La continuidad del espacio requiere entonces que todos los puntos ideales se encuentren sobre un mismo *plano ideal* que corta a cada familia de planos mutuamente paralelos a lo largo de una *recta ideal* diferente. Los fundamentalistas tiritarán de horror ante esta arbitraria multiplicación de los entes. Pero la aritmética había hecho otro tanto durante siglos, enriqueciendo poco a poco su haber inicial —formado por los números naturales, 1, 2, 3,…— con el cero, los enteros negativos, los racionales no enteros, los irracionales y los llamados “números imaginarios”.

Los puntos de una recta tienen relaciones mutuas de orden y de vecindad. Para ver cómo el punto ideal de la recta entra en este sistema de relaciones, hagamos rotar continuamente el plano Θ en torno a la recta *m* donde corta a Γ. Cuando Θ sea paralelo a *PQ* —digamos, en el instante *t*— la proyección de *Q* desde *P* sobre Θ es el punto ideal de la recta que pasa por *P* y *Q*. Inmediatamente antes de *t* dicha proyección es un punto ordinario de Θ, muy distante de *m*. Inmediatamente después de *t* la proyección vuelve a ser un punto ordinario de Θ, muy distante de *m* pero en el extremo opuesto del plano. Considerando el desplazamiento continuo de la proyección durante un breve lapso que incluya al instante *t*, concluimos que, si *A* y *B* son dos puntos cualesquiera de Θ situados uno a cada lado de la recta *m*, y la recta *a* que los une corta a *m* en *C*, el punto ideal de *a* tiene que estar entre *A* y *B* pero no se halla ni entre *A* y *C* ni entre *C* y *B*. Por lo tanto, en la geometría proyectiva los puntos de una recta están ordenados cíclicamente, como los puntos de una circunferencia. A consecuencia de esto, las relaciones de vecindad entre los puntos en el espacio proyectivo y sobre los planos proyectivos difie-

ren drásticamente de las familiares en la geometría estándar y no son nada intuitivas. Por eso, es justo decir que la geometría proyectiva entrañó una revolución del pensamiento mucho más profunda y de mayor alcance que la mera negación del postulado de Euclides.

En el nuevo ámbito, las propiedades proyectivas de las figuras pueden definirse impecablemente. Una aplicación biunívoca f del espacio proyectivo sobre sí mismo es una *colineación* si envía tres puntos colineales cualesquiera A, B y C a tres puntos $f(A)$, $f(B)$ y $f(C)$, que también son colineales. Las propiedades —y relaciones— proyectivas son aquellas que son preservadas por las colineaciones. He aquí algunos ejemplos de propiedades proyectivas: (i) de tres o más puntos, *incidir en la misma línea*, *yacer sobre el mismo plano*; (ii) de tres o más rectas, *concurrir en un mismo punto*, *yacer sobre el mismo plano*; (iii) de tres o más planos, *compartir un mismo punto*, *cortarse en la misma recta*; (iv) de una curva, *ser una cónica*; (v) de una superficie, *ser una cuádrica*.

6.3 El programa de Erlangen

En un folleto que publicó en 1872 al incorporarse como profesor a la Universidad de Erlangen, Felix Klein (1849–1925) tomó nota del enorme crecimiento y diversificación de la geometría desde 1800 y propuso un punto de vista desde el cual sus muchas ramas podían ordenarse en un sistema. Desde el punto de vista de Klein, la tarea de una rama de la geometría puede describirse así: *Dada una variedad y un grupo de transformaciones de la variedad, investigar las figuras discernibles en la variedad con respecto a aquellas propiedades que no son alteradas por las transformaciones del grupo* (Klein 1893, p. 67). En la literatura matemática de su tiempo, ‘variedad’ (*Mannigfaltigkeit* en alemán) solía designar lo que hoy llamamos ‘conjunto’, pero Klein apuntaba a algo más específico: “Dadas n variables $x_1,\ldots, x_n$, los [...] sistemas de valores que se obtienen si dejamos que las variables x tomen independientemente todos los valores reales desde $-\infty$ hasta $+\infty$ constituyen lo que llamaremos

[…] *una variedad de n dimensiones.* Cada sistema de valores particular ($x_1, \ldots, x_n$) se llama un *elemento* de la variedad” (Klein 1873, p. 116). Si *S* es una variedad en una de estas dos acepciones, llamamos una *transformación* de *S* a cualquier aplicación ☆biyectiva☆ de *S* sobre sí misma. Está claro que

(i) si T_1 y T_2 son transformaciones de *S*, entonces la aplicación compuesta $T_2 \circ T_1$, que consiste en aplicar T_2 inmediatamente después de T_1, también es una transformación de *S*;
(ii) la composición de transformaciones es asociativa, de modo que, si T_1, T_2 y T_3 son transformaciones de *S*, entonces $(T_3 \circ T_2) \circ T_1 = T_3 \circ (T_2 \circ T_1)$;
(iii) la aplicación idéntica 1_S que asigna cada elemento de *S* a sí mismo es una transformación de *S* tal que, para cualquier transformación T de *S*, se cumple que $T \circ 1_S = 1_S \circ T = T$;
(iv) para cualquier transformación T de *S* hay una transformación *inversa* que denotamos con T^{-1}, tal que $T^{-1} \circ T = 1_S$ (T^{-1} asigna a cada elemento *P* de *S* el punto al cual *P* fue asignado por T).

En virtud de las condiciones (i)–(iv), las transformaciones de *S* forman un *grupo* G_S en el sentido preciso que el término ‘grupo’ posee en el álgebra (cf. la nota 15 al capítulo 1, p. 204). G_S incluye subgrupos, esto es, subconjuntos que contienen a 1_S y satisfacen las condiciones (i) y (iv). Si *H* es un subgrupo de G_S y $\mathfrak{P}$ es una propiedad de *S* o de sus elementos o sus partes que no es afectada por las transformaciones pertenecientes a *H*, decimos que $\mathfrak{P}$ es *H*-invariante o que es ‘un invariante de *H*’. La única propiedad G_S-invariante de *S* es el número de elementos que *S* contiene (su cardinalidad $|S|$). Por otra parte, el grupo $\{1_S\}$, que consta de la identidad solamente, trivialmente preserva todas las propiedades imaginables. Entre estos dos extremos caben muchos subgrupos diferentes con toda clase de invariantes dignos de interés, dependiendo de la estructura de grupo pertinente. Si no es un conjunto arbitrario, sin estructura, sino la variedad numérica descrita por Klein, *S* hereda del cuerpo $\mathbb{R}$ de los números reales una estructura que contribuye a caracterizar los distintos subgrupos de G_S y sus invariantes respectivos. Así, el grupo de las transformaciones continuas preserva las propiedades topológicas

(relaciones de vecindad) y el grupo de las transformaciones lineales preserva las propiedades proyectivas.

¿Es posible fijar de este modo las propiedades métricas? Tradicionalmente, la distancia entre dos puntos $\langle x_1,\ldots,x_n\rangle$ y $\langle y_1,\ldots,y_n\rangle$ de una variedad numérica es igual a $\sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}$, por definición. El grupo de las isometrías comprende todas las transformaciones que preservan esta función. Sin embargo, esto no es más que una convención, adoptada para asegurar que la geometría de la variedad sea euclídea. Valiéndose de la geometría proyectiva, Klein dio con una idea mejor. Ninguna función de pares de puntos con valores reales puede ser un invariante proyectivo, pero hay una función de cuádruplos de puntos colineales que sí lo es, la llamada *razón anharmónica*. Basándose en los trabajos de Arthur Cayley (1821–1895), Klein (1871, 1873) utilizó la razón anharmónica para definir sobre regiones específicas del espacio proyectivo funciones proyectivamente invariantes que asignan distancias a pares de puntos pertenecientes a esas regiones. Notablemente, dos de las tres funciones definidas por Klein concuerdan en sus regiones respectivas con la distancia euclídea y la distancia lobachevskiana, mientras que la tercera determina otra clase de geometría no euclídea que Klein llamó *elíptica*. (En la geometría elíptica, cada recta corta todas las otras rectas coplanares con ella y los tres ángulos internos de un ángulo siempre suman más de dos rectos.) Klein llama *parabólica* a la geometría de Euclides e *hiperbólica* a la de Lobachevsky.

Veamos como el método de Klein se aplica a la geometría lobachevskiana del plano. Sea ζ una cónica real en el plano proyectivo, esto es, una cónica que solo contiene puntos reales. Sea G_ζ el conjunto de todas las colineaciones que aplican la cónica ζ sobre sí misma. G_ζ es un subgrupo del grupo proyectivo. Considérese ahora la razón anharmónica de los cuádruplos de puntos $\langle P_1,P_2,P_3,P_4\rangle$ tales que P_3 y P_4 pertenecen a ζ mientras que P_1 y P_2 recorren el interior Int(ζ) de ζ. ($P \in$ Int(ζ) si y solo si P es un punto real y ninguna recta real tangente a ζ pasa por P.) P_3 y P_4 son los puntos donde la recta determinada por P_1 y P_2 corta ζ, de modo que la razón anharmónica indicada puede considerarse como una función del par de puntos $\langle P_1,P_2\rangle$, que llamaremos f_ζ. Evidentemente f_ζ es G_ζ-invariante. De-

finimos la función $\delta : \text{Int}(\zeta) \times \text{Int}(\zeta) \to \mathbb{R}$ mediante la ecuación $\delta(P,Q) = k \log f_\zeta(P,Q)$, donde k es una constante real arbitraria diferente de 0 y $\log x$ denota el valor principal del logaritmo natural de x. Klein demostró que δ se comporta exactamente como una distancia lobachevskiana sobre Int(ζ). En otras palabras, todos los teoremas de la geometría lobachevskiana del plano valen para figuras pertinentes formadas con puntos de Int(ζ) si la distancia entre dos cualesquiera de esos puntos está dada por la función δ. Considérense, por ejemplo, cuatro puntos P_1, P_2, P_3 y P_4 de Int(ζ) tales que $\delta(P_1,P_2) = \delta(P_2,P_3) = \delta(P_3,P_4) = \delta(P_4,P_1)$. Estos cuatro puntos son los vértices de un cuadrilátero equilátero lobachevskiano $\mathfrak{C}$ que puede tener a lo sumo tres ángulos internos rectos, en cuyo caso el cuarto ángulo interno de $\mathfrak{C}$ tiene que ser agudo. (Por 'ángulo recto' se entiende aquí, como siempre, un ángulo igual a su ángulo adyacente, y dos ángulos en Int(ζ) se dicen iguales si uno es la imagen del otro por una transformación del grupo G_ζ).

Si ζ no es una cónica real ordinaria, sino una cónica de otra índole, la función δ definida como arriba se comporta en regiones apropiadas del plano proyectivo como la distancia euclídea o como la distancia de la geometría elíptica. Dependiendo, pues, de que ζ pertenezca a una u otra de tres clases de cónicas, el grupo de las colineaciones que aplican ζ sobre sí misma es estructuralmente idéntico con el grupo de isometrías del plano lobachevskiano, euclídeo o elíptico. Resultados similares se aplican al caso tridimensional, si la cónica ζ se reemplaza con una superficie cuádrica.

Este descubrimiento de Klein indujo a Bertrand Russell (1873–1970) a sostener, en su libro sobre los fundamentos de la geometría (1897), que la "forma de la externalidad" se nos manifiesta a priori en la geometría proyectiva, pero que su estructura métrica —que *solo* puede ser lobachevskiana, euclídea o elíptica— tiene que determinarse a posteriori por experimentos. Henri Poincaré (1854–1912) adoptó una posición mucho más radical: si la geometría no es más que el estudio de un grupo, "se puede decir que la verdad de la geometría de Euclides no es incompatible con la verdad de la geometría de Lobachevsky, pues la existencia de un grupo no es incompatible con la de otro grupo" (Poincaré 1887, p. 290). La aplicación a la física

es inmediata: "Entre todos los grupos posibles hemos elegido uno en particular, para referir a él todos los fenómenos físicos, tal como elegimos tres ejes de coordenadas para referir a ellos una figura geométrica" (Ibid., p. 291). La elección de este grupo particular está motivada por su simplicidad matemática, pero también por el hecho de que "existen en la naturaleza ciertos cuerpos notables que se llaman *sólidos* y la experiencia nos enseña que los diversos movimientos posibles de estos cuerpos se relacionan entre sí aproximadamente del mismo modo que las diferentes operaciones del grupo escogido" (Ibid.).

El punto de vista presentado por Klein en el programa de Erlangen gozó de gran favor entre los matemáticos y los filósofos. Alcanzó un gran éxito cuando Minkowski (1909) mostró que la espina dorsal de la teoría especial de la relatividad de Einstein era la geometría (espaciotemporal) del grupo de Lorentz. Klein (1911) subraya con toda razón este triunfo de sus ideas. Pero el programa de Erlangen no era capaz de cubrir la geometría diferencial de Riemann (§6.5) que Einstein (1915*i*, 1916*e*) colocó en el centro de su teoría general de la relatividad.

6.4 Perfeccionamiento de la axiomática

Según Aristóteles, el conocimiento científico (ἐπιστήμη) tiene que expresarse en proposiciones que se deducen de una lista finita de enunciados de suyo evidentes (axiomas) y solo emplean términos definidos a partir de una lista finita de términos que se entienden por sí mismos (primitivos). Durante más de dos mil años se dio por descontado que los *Elementos* de Euclides efectivamente satisfacen estas exigencias de Aristóteles. De hecho, empero, hay una laguna lógica ya en Euclides I.1 —"Construir un triángulo equilátero sobre un segmento recto dado"—, pues la solución de este problema reposa sobre un tácito supuesto de continuidad. Tampoco es claro que Euclides

creyera que sus postulados son evidentes de suyo; más bien sugiere lo contrario cuando los llama αἰτήματα, esto es, *peticiones*.

La idea de asegurar el conocimiento mediante la deducción lógica a partir de premisas incuestionables fascinó a científicos modernos como Galileo y Newton, que practicaron gustosamente la axiomática, al menos como género literario, lo mismo que Spinoza en su *Ética*. Sin embargo, un ejemplo de axiomatización satisfactoria y responsable de una rama del conocimiento no se produjo hasta 1882, cuando Moritz Pasch (1843–1930) publicó sus *Lecciones de geometría moderna*.

Pasch veía a la geometría como una ciencia natural, cuya provechosa utilización por otras ciencias y en la vida práctica se basa "exclusivamente en el hecho de que originalmente los conceptos geométricos concuerdan exactamente con los objetos empíricos" (Pasch 1882, p. iii). La geometría se distingue de otras ciencias naturales porque obtiene muy pocos conceptos y leyes directamente de la experiencia y aspira a derivar de ellos las leyes de fenómenos más complejos por medios puramente deductivos. Pasch encapsula el fundamento empírico de la geometría en un núcleo de conceptos básicos y proposiciones básicas o axiomas. Los conceptos básicos se refieren a la forma, tamaño y posición relativa de los cuerpos. No se los define, pues ninguna definición podría sustituir la "exhibición de los objetos naturales apropiados", que es la única vía para entender nociones tan simples e irreductibles (ibid., p. 16). Todos los demás conceptos geométricos tienen que definirse en último término a partir de los básicos. Los conceptos básicos están conectados entre sí por los axiomas, que "enuncian lo que se observa en ciertos diagramas muy simples" (p. 43). Todas las demás proposiciones geométricas deben demostrarse a partir de los axiomas mediante el más riguroso método deductivo. Todo lo que se necesita para demostrarlos tiene que estar consignado, sin excepción, en los axiomas. Estos deben, pues, recoger todo el material empírico elaborado por la geometría, de modo que "una vez que se los haya establecido no haga falta recurrir nuevamente a la percepción sensible" (p. 17). "Toda conclusión que figure en una prueba tiene que hallar su confirmación en el diagrama, pero no es justificada por el diagrama, sino por cierta proposición (o

definición) precedente" (p. 43). Pasch entendió claramente las implicaciones de su método. Escribe:

> Para que la geometría sea verdaderamente deductiva, el proceso de inferencia debe ser en todas sus partes independiente del *significado* de los conceptos geométricos, tal como debe ser independiente de los diagramas. Todo lo que hay que considerar son las *relaciones* entre los conceptos geométricos, consignadas en las proposiciones y definiciones. En el curso de la deducción es lícito y útil tener presente el significado de los conceptos geométricos que figuran en ella, pero *no es en absoluto necesario*. Más bien, si resulta efectivamente necesario, ello muestra que hay una laguna en la demostración y, si la laguna no puede eliminarse modificando el razonamiento, que las premisas son demasiado débiles para sostenerlo.
>
> (Pasch 1882, p. 98)

Las *Lecciones de geometría moderna* de Pasch se refieren a la geometría proyectiva. La primera axiomatización de la geometría euclídea que cumple las exigencias de Pasch apareció en 1899 en los *Fundamentos de la geometría* de David Hilbert (1862–1943), una obra que tuvo una influencia enorme sobre las matemáticas y la filosofía del siglo xx. Hilbert invita al lector a considerar tres colecciones de objetos cualesquiera, que llama 'puntos', 'rectas' y 'planos', y cinco relaciones indefinidas, a saber, (i) la relación de *incidencia* entre un punto y una recta, (ii) la relación de *incidencia* entre una recta y un plano, (iii) la relación *estar entre* de un punto con otros dos, (iv) la relación de *congruencia* entre dos *segmentos* (esto es, entre dos pares de puntos) y (v) la relación de *congruencia* entre dos *ángulos* (esto es, entre dos clases de equivalencia formadas por triples de puntos). Las condiciones prescritas en los veinte axiomas de Hilbert —incluyendo el Axioma de Completud que agregó en la segunda edición— son suficientes para caracterizar hasta un isomorfismo esos objetos y relaciones. Pero un isomorfismo, claro está, puede existir entre sistemas muy diferentes de objetos, que intuitivamente tengan muy poco que ver entre sí. Hilbert utilizó esta propiedad de las teorías axiomáticas para investigar la independencia de algunos de sus axiomas. Para probar la independencia de un determinado axioma *A* pro-

puso realizaciones (“modelos”) de la estructura determinada por la unión de la negación de *A* y el resto de sus axiomas. Gottlob Frege (1848-1925) objetó que los axiomas geométricos preservados en estos ejercicios solo podían aplicarse a los extravagantes modelos propuestos por Hilbert si uno adulteraba el sentido natural de las palabras. El 29 de diciembre de 1899, Hilbert le respondió así:[7]

> Cada teoría es solo un tinglado o esquema de conceptos junto con las relaciones necesarias entre ellos, y los elementos básicos pueden concebirse como a uno le dé la gana. Si tomo como mis puntos cualquier sistema de cosas, por ejemplo, el sistema *amor*, *ley*, *deshollinador*, […] y postulo todos mis axiomas como relaciones entre estas cosas, mis teoremas —por ejemplo, el de Pitágoras— también valen para estas cosas […]. Esta propiedad de las teorías nunca puede ser un inconveniente y en todo caso es inevitable.

Todo esto se desprende, por cierto, de la naturaleza misma de la axiomática, como se explica en el pasaje de Pasch citado en la p. 151. La existencia de permutaciones semánticas que preservan la verdad no era ninguna novedad en geometría, después que Gergonne (1771-1859) llamó la atención en 1825 hacia el siguiente *principio de dualidad*: En la geometría proyectiva plana, cualquier proposición verdadera da lugar a otra, igualmente verdadera, que se obtiene reemplazando sistemáticamente en la primera ‘punto’ con ‘recta’, ‘colineal’ con ‘concurrente’, ‘une’ con ‘corta’, y viceversa. (En la geometría proyectiva del espacio, hay dualidad entre puntos y planos.) El mismo resultado se obtiene, obviamente, si en vez de las palabras, se intercambian los respectivos significados.

6.5 La geometría diferencial de Riemann

En una conferencia "Sobre las hipótesis que están en la base de la geometría", leída ante la Facultad de Filosofía de la Universidad de Göttingen en 1854 y publicada póstumamente en 1867, Bernhard Riemann (1826–1866) presentó algunas ideas radicalmente innovadoras sobre esta materia. Señaló que las propiedades mensurables de un conjunto o variedad discreta pueden determinarse contando. (Piénsese en la población de un país, o en propiedades como el ingreso total, mediano o per cápita, o la proporción de católicos o de analfabetos.) Pero las propiedades mensurables de una variedad continua no pueden determinarse de este modo. En particular, las propiedades mensurables del espacio físico, que son el tema de la geometría, dependen de las fuerzas enlazantes que actúan sobre él.

La distancia entre dos puntos del espacio se puede establecer mediante una vara o una cinta de medir, o con instrumentos ópticos, y el resultado depende esencialmente del comportamiento físico de los medios utilizados. Hasta ese momento, las propiedades mensurables del espacio se habían descrito con éxito conforme a la geometría euclídea. Sin embargo, "los conceptos empíricos en que se basan las determinaciones métricas del espacio —los conceptos de cuerpo rígido y de rayo de luz— pierden su validez en lo infinitamente pequeño; es lícito pensar, pues, que las relaciones métricas del espacio en lo infinitamente pequeño no concuerdan con los supuestos de la geometría, y de hecho tendríamos que aceptarlo si ello permitiese explicar los fenómenos de un modo más simple" (Riemann 1854, p. 149). Para preparar a los físicos para esta posibilidad, Riemann propuso una concepción más general de la geometría. Su esquema básico admite un grado de generalidad mucho mayor que el que alcanza en efecto; pero, a juicio suyo, por ahora es suficiente caracterizar la geometría de las variedades continuas de modo que concuerde óptimamente con la geometría euclídea en un pequeño entorno de cada punto.

Riemann extiende a *n* dimensiones los métodos empleados por Gauss (1828) en sus investigaciones acerca de la geometría intrínseca de las superficies curvas discernibles en el espacio euclídeo. (Se la

llama "intrínseca" porque considera las propiedades métricas que pertenecen a las superficies mismas, independientemente del modo como yacen o están incrustadas en el espacio—por ejemplo, el teorema de Pitágoras vale con exactamente la misma precisión sobre una hoja de papel bien estirada, o enrollada como cucurucho, o arrugada como pelota, siempre que al deformarla la hoja no se dilate, ni desgarre, ni contraiga). Mirando hacia atrás a la obra de Gauss, se gana una mejor comprensión intuitiva de los conceptos de Riemann.[8] Pero, en aras de la concisión y la claridad es preferible mirar hacia adelante y valerse de los conceptos introducidos por matemáticos posteriores para entender la propuesta de Riemann.

Los continuos multidimensionales con que Riemann se ocupa son ejemplos de lo que ahora llamamos una variedad diferenciable *n*-dimensional real. En la §7.8 doy una caracterización breve y precisa de su estructura. Luego, en las §7.12 y §7.13 explico las nociones de métrica riemanniana, geodésica y curvatura. Una *métrica riemanniana* asigna longitudes a las curvas trazadas en una variedad *n*-dimensional. En esto se basa el concepto de *geodésica* o curva de longitud extrema: la más corta o la más larga de todas las curvas que recorren caminos vecinos entre dos puntos dados. En su investigación de las superficies curvas, Gauss introdujo una función con valores reales, la *curvatura gaussiana,* que evalúa en cada punto de la superficie, en términos de su geometría intrínseca, la medida en que esta se aparta de la geometría plana en ese punto y su vecindario. Riemann extendió este concepto de curvatura a las variedades *n*-dimensionales provistas de una métrica del tipo introducido por él. Valiéndose de este concepto de curvatura, Riemann pudo caracterizar con gran elegancia las variedades de esta clase en que todas las figuras pueden moverse libremente de una parte a otra sin cambiar de forma ni tamaño. Son las variedades riemannianas de curvatura constante. Esta idea se deja combinar muy bien con la clasificación de las geometrías métricas debida a Klein. Si los consideramos como variedades riemannianas tridimensionales, el espacio euclídeo tiene curvatura cero, el espacio lobachevskiano tiene curvatura constante negativa y el espacio elíptico tiene curvatura constante positiva. Conforme al programa de Erlangen, cada una de estas geometrías de curvatura constante se

caracteriza por su propio grupo de isometrías. Pero la concepción de Klein es demasiado estrecha para abarcar a todas las geometrías riemannianas. En el caso general en que la curvatura varía de punto en punto, el grupo de isometrías de una variedad riemanniana n-dimensional es el grupo trivial que consta de la identidad solamente. La estructura de este grupo obviamente no suministra ninguna información sobre la respectiva geometría.[9]

7

Apéndice matemático

7.1 Conjuntos y aplicaciones

Sea S un conjunto. Escribimos $x \in S$ para indicar que x es un elemento de S, y $x \notin S$ para indicar que x no es un elemento de S. Para referirnos al conjunto de todos los objetos x que tienen la propiedad P usamos la expresión '$\{x : x$ tiene la propiedad $P\}$'. Un conjunto R es un *subconjunto* de S (simbolizado: $R \subseteq S$) si todo elemento de R es un elemento de S. Según esta definición, todo conjunto es un subconjunto de sí mismo. Además, el conjunto vacío $\varnothing = \{x : x \neq x\}$ es un subconjunto de cada conjunto. El *complemento* de R en S (simbolizado: $S \setminus R$) es el conjunto de todos los elementos de S que no son elementos de R, esto es, el conjunto $\{x : x \in S \;\&\; x \notin R\}$. El conjunto de todos los subconjuntos de un conjunto S se llama el *conjunto potencia* de S y se designa con $\mathcal{P}S$. La teoría de conjuntos postula que, dado un conjunto cualquiera S, también está dado el conjunto $\mathcal{P}S$. Si A y B son conjuntos, el *producto cartesiano* $A \times B$ es el conjunto de todos los pares ordenados $\langle a,b \rangle$, tales que $a \in A$ y $b \in B$. Escribimos A^2 en vez de $A \times A$, A^3 en vez de $(A \times A) \times A$, y A^n en vez de $A^{n-1} \times A$.[1]

Si A y B son conjuntos, una *aplicación* f de A en B es una correspondencia que asigna a cada objeto $x \in A$ un y solo un elemento de B, designado con $f(x)$. A es el *dominio* de f, ocasionalmente llamado dom f. B es el *codominio* de f. $f(x)$ es el *valor* que f asigna al *argumento* x. La expresión 'f: $A \to B$' denota la aplicación f con dominio A y codominio B. La expresión 'f: $x \mapsto f(x)$' denota la aplicación f que a cada argumento x asigna el valor $f(x)$. (Por ejemplo, f: $\mathbb{R} \to \mathbb{R}$; $x \mapsto x^2$ es la aplicación que a cada número real x asigna su cuadrado x^2.) El *recorrido* de f —también llamado la *imagen* de A por f— es el conjunto de los valores de f, esto es, el con-

junto $\{x : x \in B \ \& \ x = f(y)$ para algún $y \in A\}$; lo designaremos con $f[A]$. Si D es un subconjunto no vacío de $f[A]$, entonces el conjunto $\{x : x \in A \ \& \ f(x) \in D\}$ es la *imagen inversa* de D por f, simbolizada $f^{-1}[D]$. Si $D = \{y\}$ —un conjunto unario que contiene el solo elemento y— su imagen inversa por f se simboliza simplemente asi: $f^{-1}(y)$. f es una aplicación *inyectiva* o un *inyección* si, para cualesquiera elementos x, $y \in$ dom f, $f(x) = f(y)$ implica que $x = y$. f se dice *epiyectiva* si aplica el dominio A *sobre* el codominio B, esto es, si cada elemento de B es asignado por f a algún elemento de A. f es una aplicación *biyectiva* o una *biyección* si f es a la vez inyectiva y epiyectiva. Si f: $A \to B$ es inyectiva, hay una aplicación $f^{-1} : f[\mathrm{A}] \to A$, que asigna a cada valor de f el argumento respectivo. f^{-1} es la *inversa* de f; es obviamente biyectiva. Si $M \subseteq A =$ dom f, la aplicación que asigna a cada $x \in M$ el valor $f(x)$ es la *restricción* de f a A. Normalmente, se la denota también con f, pero cuando es necesario utilizar una notación más precisa, la restricción de una aplicación f a una parte A de dom f se simboliza $f \restriction A$.

Definiré ahora la aplicación $g \circ f$, compuesta de la aplicación f por la aplicación g. Si el recorrido de f es un subconjunto del dominio de g, $g \circ f$ asigna a cada $x \in$ dom f el valor de g en $f(x)$. Si $f[\text{dom } f] \nsubseteq$ dom g, pero $f[\text{dom } f] \cap \text{dom } g = U \neq \varnothing$, la expresión $g \circ f$ designa la aplicación compuesta de la restricción de f a $f^{-1}[U]$ por la restricción de g a U.

Una aplicación cuyos valores son números (naturales, enteros, racionales, reales o complejos) se llama normalmente una *función.*

7.2 Espacios topológicos

Sea S un conjunto cualquiera. Una *topología* en S es una colección de T de subconjuntos de S, llamados *abiertos*, tales que

(i) S es un abierto;
(ii) el conjunto vacío $\varnothing$ es un abierto;
(iii) si A y B son abiertos, su intersección $A \cap B$ es un abierto;
(iv) la unión de cualquier colección finita o numerable de abiertos es un abierto.

La estructura $\langle S,T\rangle$ es un *espacio topológico*; a menos que ello genere ambigüedad, lo designamos por el nombre del conjunto subyacente S. Los elementos de un espacio topológico S se llaman *puntos*. Si T_1 y T_2 son dos topologías diferentes en el mismo conjunto y $T_1 \subseteq T_2$ se dice que T_1 es *más gruesa* que T_2 y T_2 *mas fina* que T_1. Si A es un abierto de $\langle S,T\rangle$, su complemento $S \setminus A$ se dice *cerrado*.

Si $A \subseteq S$, la *topología de subconjunto* en A se define así: la intersección de A con cualquier abierto de S es un abierto de A. Un subconjunto de S provisto de la correspondiente topología de subconjunto es un *subespacio* de S. De aquí en adelante, consideramos a cada subconjunto de un espacio topológico S como un subespacio de S en este sentido.

Sea p un punto del espacio topológico S. Un *entorno* de p es cualquier subconjunto de S que incluya un abierto que contiene a p.

Si S y T son espacios topológicos, la aplicación f: $S \to T$ se dice *continua* si la imagen inversa $f^{-1}[A]$ de cualquier abierto A de T es un abierto de S. Un *homeomorfismo* es una aplicación biyectiva continua cuya inversa también es continua. Dos espacios topológicos S y T son *homeomorfos* si hay un homeomorfismo f: $S \to T$.

Utilizando conceptos topológicos se puede caracterizar la finitud de un espacio sin apelar a nociones métricas (distancia, volumen). Sea $\langle S,T\rangle$ un espacio topológico. Una colección de abiertos $R \subseteq T$ es un *recubrimiento abierto* de S si todo punto de S pertenece por lo menos a uno de los abiertos de R. $\langle S,T\rangle$ es un espacio topológico *compacto* si y solo si todo recubrimiento abierto de S incluye un recubrimiento abierto *finito*, esto es, si y solo si todo recubrimiento $R \subseteq T$ incluye un recubrimiento $R' \subseteq R$ tal que R' contiene solo un número finito de abiertos.

7.3 El espacio topológico $\mathbb{R}$

Supongo que el lector está familiarizado con el cuerpo $\mathbb{R}$ de los números reales. La topología estándar de $\mathbb{R}$ se define por la condición siguiente: si a y b son dos números reales cualesquiera, tales que $a < b$, el intervalo abierto $\{x : a < x < b\}$ es un abierto de $\mathbb{R}$; este

intervalo se denota habitualmente con (a,b). Obsérvese que la función f: $(a,b) \rightarrow (0,1)$ definida por la ecuación siguiente es un homeomorfismo:

$$f(x) = \frac{x-a}{b-a} \tag{1}$$

Por otra parte, la siguiente función g aplica homeomórficamente el intervalo (0,1) sobre $\mathbb{R}$:

$$g(x) = \exp\left(\frac{x}{1-x}\right) - \frac{1}{x} \tag{2}$$

Por lo tanto, la función compuesta $g \circ f$ establece un homeomorfismo entre un abierto cualquiera (a,b) y todo el espacio $\mathbb{R}$.

7.4 *Espacios vectoriales*

En esta sección definiré el concepto de *espacio vectorial real n-dimensional*. Sea V un conjunto cualquiera. Consideremos una aplicación de $V \times V$ en V que asigna a cada par $\langle u,v \rangle \in V \times V$, el elemento $a + b \in V$. Supongamos que

(i) para cualesquiera u, v, $w \in V$, $u + (v + w) = (u + v) + w$,
(ii) hay un elemento $\mathbf{0} \in V$ tal que, para todo $w \in V$, $w + \mathbf{0} = w$,
(iii) para cada $w \in V$ hay un elemento $w^- \in V$ tal que $w + w^- = \mathbf{0}$,
(iv) para cualesquiera u, $v \in V$, $u + v = v + u$.

La estructura $\mathcal{V} = \langle V,+,\mathbf{0} \rangle$ es entonces un *grupo abeliano* (cf. la definición de *grupo* en las pp. 204–205). Supongamos ahora que hay una aplicación φ: $\mathbb{R} \times V \rightarrow V$ que se combina con la operación $+$ en $\mathcal{V}$ y con la multiplicación ordinaria en $\mathbb{R}$ de acuerdo con las reglas **V1–V2** que enunciaré enseguida. Entonces decimos que $\mathcal{V}$ es un *espacio vectorial real* o *espacio vectorial* sobre $\mathbb{R}$. En tal caso, los elementos de V se llaman *vectores*; los números reales se llaman *escalares*; la operación $+$, *suma vectorial*, y la aplicación φ, *multiplicación por escalares*. Sean, pues, a y b cualesquiera escalares;

v y **w**, cualesquiera vectores. En vez de $\varphi(a,\mathbf{v})$, escribimos $a\mathbf{v}$. Las reglas son estas:

V1 $a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w}$.
V2 $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$.
V3 $a(b\mathbf{v}) = (ab)\mathbf{v}$.
V4 $1\mathbf{v} = \mathbf{v}$ (donde 1 es, por cierto, el número $1 \in \mathbb{R}$)

Consideremos un conjunto no vacío $W \subseteq V$, cerrado bajo la suma vectorial y la multiplicación por escalares (esto significa que, si $\mathbf{v}$, $\mathbf{w} \in W$ y $a \in K$, $\mathbf{v} + \mathbf{w}$ y $a\mathbf{v}$ pertenecen a W). Denotando con $+\restriction W$ la restricción a W de la suma vectorial en $\mathcal{V}$, tenemos que $\mathcal{W} = \langle W,+\restriction W,\mathbf{0}\rangle$ es un espacio vectorial real; se dice que $\mathcal{W}$ es un *subespacio* de $\mathcal{V}$. Normalmente, por cierto, se escribe + en vez de $+\restriction W$.

Si $\{\mathbf{v}_1, \mathbf{v}_2,\ldots, \mathbf{v}_n\}$ es un conjunto de n vectores de $\mathcal{V}$ y $a_1, a_2,\ldots, a_n$ es una lista de escalares, el vector $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 +\ldots+ a_n\mathbf{v}_n$ es una *combinación lineal* de los vectores $\mathbf{v}_1, \mathbf{v}_2,\ldots, \mathbf{v}_n$. Obviamente, el conjunto de todas las combinaciones lineales posibles de un conjunto no vacío de vectores $\{\mathbf{v}_1, \mathbf{v}_2,\ldots, \mathbf{v}_n\}$ es un subespacio de $\mathcal{V}$, el subespacio *generado* por ese conjunto. Se dice que un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2,\ldots, \mathbf{v}_n\}$ es *linealmente independiente* si y solo si la combinación lineal $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 +\ldots+ a_n\mathbf{v}_n = \mathbf{0}$ únicamente en el caso de que todos los escalares $a_1 = a_2 =\ldots= a_n = 0$. En tal caso, obviamente, ninguno de los vectores de ese conjunto es igual a una combinación lineal de los demás.

Sea $B = \{\mathbf{v}_1, \mathbf{v}_2,\ldots, \mathbf{v}_n\}$ un conjunto finito linealmente independiente de vectores de $\mathcal{V}$. B es una *base* de $\mathcal{V}$ si todo vector $\mathbf{v} \in V$ es igual a una combinación lineal de vectores de B: $\mathbf{v} = \sum_{i=1}^{n} a_i\mathbf{v}_i$. En tal caso, los escalares a_i son los *componentes* de **v** *relativos a la base B*. Obsérvese que, si $\mathcal{V}$ tiene una base formada por n vectores, ningún conjunto linealmente independiente de vectores de $\mathcal{V}$ puede contener más de n vectores. Por lo tanto, toda base de $\mathcal{V}$ contiene exactamente n vectores. En tal caso, se dice que $\mathcal{V}$ es un espacio vectorial n-dimensional. Si $\mathcal{V}$ y $\mathcal{W}$ son dos espacios vectoriales reales, la aplicación $f: \mathcal{V} \rightarrow \mathcal{W}$ se dice *lineal* si preserva la suma vectorial y la multiplicación por escalares, es decir, si $f(a\mathbf{v} + b\mathbf{w}) = af(\mathbf{v}) + bf(\mathbf{w})$,

cualesquiera que sean **v**, **w** $\in$ V y a, $b \in \mathbb{R}$. En particular, si $\mathcal{W} = \mathbb{R}$, se dice que f es una *función lineal* sobre $\mathcal{W}$.

El conjunto de todas las funciones lineales sobre el espacio vectorial real n-dimensional $\mathcal{V}$ tiene también la estructura de un espacio vectorial real n-dimensional; lo llamamos el *espacio dual* de $\mathcal{V}$, y lo designamos con $\mathcal{V}^*$. Si φ y ψ pertenecen a $\mathcal{V}^*$, la suma $\varphi + \psi$ se define por la condición: $(\varphi + \psi)(\mathbf{v}) = \varphi(\mathbf{v}) + \psi(\mathbf{v})$ para cada $\mathbf{v} \in \mathcal{V}$; es claro que $\varphi + \psi$ pertenece a $\mathcal{V}^*$, pues, si a, $b \in \mathbb{R}$ y $\mathbf{v}$, $\mathbf{w} \in \mathcal{V}$, $(\varphi + \psi)(a\mathbf{v}+b\mathbf{w}) = \varphi(a\mathbf{v}+b\mathbf{w}) + \psi(a\mathbf{v}+b\mathbf{w}) = a\varphi(\mathbf{v}) + b\varphi(\mathbf{w}) + a\psi(\mathbf{v}) + b\psi(\mathbf{w}) = a(\varphi(\mathbf{v})+\psi(\mathbf{v})) + b(\varphi(\mathbf{w}) +\psi(\mathbf{w})) = a(\varphi + \psi)(\mathbf{v}) + (\varphi + \psi)(\mathbf{w})$. El producto $a\varphi$ de $\varphi \in \mathcal{V}^*$ por $a \in \mathbb{R}$ se define por la condición: $(a\varphi)(\mathbf{v}) = a\varphi(\mathbf{v})$ para cada $\mathbf{v} \in \mathcal{V}$; es fácil ver que $a\varphi \in \mathcal{V}^*$. Si$\{\mathbf{v}_1,\ldots, \mathbf{v}_n\}$ es una base de $\mathcal{V}$, la *base dual* $\{\mathbf{v}_1^*,\ldots, \mathbf{v}_n^*\}$ de $\mathcal{V}^*$ se define así: $\mathbf{v}_k^*$ es la función lineal que a cada $\mathbf{v}_j$ asigna el valor δ^k_j, esto es, 1 si $k = j$ y 0 si $k \neq j$ $(1 \leq j,k \leq n)$. La correspondencia $\mathbf{v}_k \mapsto \mathbf{v}_k^*$ determina un isomorfismo $\mathcal{V} \to \mathcal{V}^*$, el cual depende de la selección arbitraria de una base. En cambio, entre $\mathcal{V}$ y $\mathcal{V}^{**}$ —el espacio dual de $\mathcal{V}^*$— hay un isomorfismo que no depende de una base particular. Este isomorfismo canónico asigna a cada $\mathbf{v} \in \mathcal{V}$ la función lineal $\mathbf{v}^{**} : \mathcal{V}^* \to \mathbb{K}$ definida por la condición $\mathbf{v}^{**}(f) = f(\mathbf{v})$ para cada $f \in \mathcal{V}^*$. En la literatura matemática, **v** se identifica con $\mathbf{v}^{**}$ y no se distingue entre $\mathcal{V}$ y $\mathcal{V}^{**}$.

El concepto de *función multilineal* se define de un modo similar. Sean $\mathcal{V}_1,\ldots,\mathcal{V}_q$ espacios vectoriales de igual número de dimensiones. La función $f: \mathcal{V}_1 \times \ldots \times \mathcal{V}_q \to \mathbb{R}$ es *q-lineal* si es lineal en cada índice r $(1 \leq r \leq q)$, esto es, si para todo r y todo $\mathbf{u}_s \in \mathcal{V}_\mathbf{s}$ $(s \neq r)$ se cumple que, cualesquiera que sean a, $b \in \mathbb{R}$ y $\mathbf{v}$, $\mathbf{w} \in \mathcal{V}_r$,

$$f(\mathbf{u}_1,\ldots,\mathbf{u}_{r-1},a\mathbf{v}+b\mathbf{w},\mathbf{u}_{r+1},\ldots,\mathbf{u}_q) = \\ af(\mathbf{u}_1,\ldots,\mathbf{u}_{r-1}, \mathbf{v},\mathbf{u}_{r+1},\ldots,\mathbf{u}_{r-1}) + bf(\mathbf{u}_1,\ldots,\mathbf{u}_{r-1}, \mathbf{w},\mathbf{u}_{r+1},\ldots,\mathbf{u}_{r-1})$$

Si $\mathcal{V}_1 = \ldots = \mathcal{V}_q = \mathcal{V}$ se dice que f es una función q-lineal *sobre* $\mathcal{V}$.

Una función bilineal $f: \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ es un *producto escalar* sobre $\mathcal{V}$ si, para cualesquiera vectores $\mathbf{v}$, $\mathbf{w} \in \mathcal{V}$, $f(\mathbf{v},\mathbf{w}) = f(\mathbf{w},\mathbf{v})$. Si f además es positivo definido, esto es, si para cada $\mathbf{v} \in \mathcal{V}$ $f(\mathbf{v},\mathbf{v}) = 0$ si y solo si $\mathbf{v} = 0$, y de otro modo, $f(\mathbf{v},\mathbf{v}) > 0$, decimos que f es un *producto interno* sobre $\mathcal{V}$. Normalmente, se escribe $\mathbf{v} \cdot \mathbf{w}$ en vez de $f(\mathbf{v},\mathbf{w})$;

adviértase que los valores de f son escalares, a diferencia del llamado *producto vectorial* que asigna a $\mathbf{v}$, $\mathbf{w} \in \mathcal{V}$ un vector $\mathbf{v} \times \mathbf{w} \in \mathcal{V}$.

Si los espacios vectoriales $\mathcal{V}$ y $\mathcal{W}$ son ambos n-dimensionales siempre hay una biyección lineal entre ellos. Esto implica que todos los espacios vectoriales del mismo número de dimensiones y sobre el mismo cuerpo son isomorfos.

Si ninguna colección finita de vectores de $\mathcal{V}$ es una base de $\mathcal{V}$, se dice que $\mathcal{V}$ tiene infinitas dimensiones.

El concepto de espacio vectorial complejo n-dimensional se define de un modo similar, reemplazando $\mathbb{R}$ con $\mathbb{C}$.[2] En la física actual tiene mucha importancia también el concepto de espacio vectorial complejo de infinitas dimensiones; pero no hace falta explicarlo aquí, pues no juega ningún papel en la teoría de la relatividad.

7.5 El espacio vectorial topológico $\mathbb{R}^n$

$\mathbb{R}^n$ es el conjunto de todos los n-tuplos $\langle x_1,\ldots,x_n\rangle$ de elementos de $\mathbb{R}$. La topología estándar de $\mathbb{R}^n$ es la más gruesa que cumple la condición siguiente: si $x_1,\ldots, x_n$ y r son cualesquiera números reales, la esfera abierta $\{\langle x_1,\ldots,x_n\rangle : x_1{}^2 + \ldots + x_n{}^2 < r^2\}$ es un abierto. $\mathbb{R}^n$ es además un espacio vectorial sobre $\mathbb{R}$ si (i) la suma de vectores y (ii) la multiplicación por escalares se definen así:

(i) $\langle x_1,\ldots,x_n\rangle + \langle y_1,\ldots,y_n\rangle = \langle x_1 + y_1,\ldots,x_n + y_n\rangle$ y
(ii) $r(\langle x_1,\ldots,x_n\rangle) = \langle rx_1,\ldots,rx_n\rangle$

para cualesquiera números reales $x_1,\ldots, x_n$, $y_1,\ldots, y_n$, r.

Presupongo que el lector tiene una familiaridad mínima con el cálculo diferencial e integral de funciones de una o más variables reales y está dispuesto a aceptar, sin más explicaciones, que el cálculo puede extenderse a aplicaciones cuyo dominio es un abierto de $\mathbb{R}^n$ y cuyo codominio es un abierto de $\mathbb{R}^m$, para cualesquiera enteros positivos n y m (cf. Spivak 1965). Sea f una aplicación de este género. Digo que f es *lisa en un punto* p si f está definida en p y, cualquiera que sea el entero positivo r y el índice k ($1 \leq k \leq n$), la derivada parcial de

r-ésimo orden $\partial^r f / \partial x_k^r$ existe y es continua en p. f es *lisa* (sin más) si es lisa en cada punto de su dominio.[3]

7.6 Variedades topológicas

Sea S un espacio topológico. Supóngase que cada punto p de S tiene un entorno homeomorfo a $\mathbb{R}^n$. Entonces S es una *variedad topológica n*-dimensional. Supongamos, por ejemplo, que S es superficie de un huevo exenta de toda aspereza, de modo que sea perfectamente lisa hasta en las regiones más pequeñas. Fotografiemos el huevo desde todos lados, asegurándonos de que cada punto de la superficie aparece retratado por lo menos una vez. La superficie se divide así en sectores que se traslapan, cada uno de los cuales está representado en una fotografía. Apliquemos ahora cada sector X biyectivamente en un abierto de $\mathbb{R}^2$ asignando a cada punto p en el interior de X el par de números $\langle x^1(p), x^2(p) \rangle$, donde $x^1(p)$ es la distancia en metros desde el punto que representa a p en la fotografía de X hasta el borde izquierdo de esa foto y $x^2(p)$ es la distancia en metros desde ese mismo punto hasta el borde inferior de la fotografía. Supongamos ahora que S está provisto de una topología en virtud de la cual todas las aplicaciones así definidas son homeomorfismos de su dominio sobre el recorrido respectivo. El lector debe comprobar que, con esta topología, S es una variedad topológica de 2-dimensiones de acuerdo con la definición precedente. Es claro, por otra parte, que S *no es homeomórfica a* $\mathbb{R}^2$. (En efecto: si p es un punto sobre el huevo, hay por lo menos una aplicación x que le asigna a p un par de números reales $\langle x^1(p), x^2(p) \rangle = x(p)$; sin embargo, se necesitan por lo menos dos aplicaciones para cubrir el huevo entero).

7.7 Cartas y atlas

Sea S una variedad topológica n-dimensional. Una *carta* de S es un homeomorfismo de un abierto de S en un abierto de $\mathbb{R}^n$. En virtud de esta definición, una carta f asigna a cada punto p de su dominio U_f un n-tuplo de números reales, $\langle f^1(p),\ldots,f^n(p)\rangle$; el n-ésimo número $f^k(p)$ es la k-ésima *coordenada* del punto p por la carta f; la función $f^k\colon U_f \to \mathbb{R}$ es la k-ésima *coordenada* de la carta f ($1 \leq k \leq n$). En la literatura más antigua —y aún hoy en muchas obras de física— la aplicación f se llama *sistema de coordenadas* (en vez de *carta*). Si f y g son dos cartas de S definidas, respectivamente, en los abiertos U_f y U_g, la aplicación compuesta $f \circ g^{-1}$ se llama la *transformación de coordenadas* entre g y f. Si $U_f \cap U_g = \varnothing$, esta aplicación es trivial. Pero si U_f se traslapa con U_g la transformación de coordenadas $f \circ g^{-1}$ y su inversa $g \circ f^{-1}$ son homeomorfismos de un abierto $\mathbb{R}^n$ sobre otro; $f \circ g^{-1}$ aplica $g[U_f \cap U_g]$ sobre $f[U_f \cap U_g]$, y $g \circ f^{-1}$ aplica $f[U_f \cap U_g]$ sobre $g[U_f \cap U_g]$.

Un *atlas* de S es una colección de cartas de S tal que cada punto de S pertenece al dominio de por lo menos una de las cartas de la colección.

7.8 Variedades diferenciables

Supongamos ahora que nuestra variedad topológica n-dimensional S está provista de un atlas $\mathcal{A}$ con la siguiente propiedad: Todas las transformaciones de coordenadas entre cartas de $\mathcal{A}$ son lisas en todos los puntos donde están definidas. (Este requisito tiene sentido, pues cada transformación entre cartas de $\mathcal{A}$ aplica un abierto de $\mathbb{R}^n$ en $\mathbb{R}^n$.) En tal caso, digo que S, equipado con el atlas $\mathcal{A}$, es una *variedad diferenciable n-dimensional* o *variedad n-dimensional*.

Este concepto también se puede definir directamente, en vez de basarlo en el concepto de variedad topológica. Tomemos un conjunto cualquiera no vacío S. Llamo n-carta de S a una aplicación biyectiva de S en un abierto $\mathbb{R}^n$. Un n-atlas diferenciable de S es una colección

de n-cartas de S tal que (i) todo punto de S está en el dominio de por lo menos una de ellas y (ii) toda transformación de coordenadas entre cartas de la colección es lisa dondequiera esté definida.[4] El conjunto S provisto de un n-atlas diferenciable $\mathscr{A}$ es una *variedad diferenciable real n-dimensional* o, más brevemente, una *variedad n-dimensional*. La variedad n-dimensional $\langle S, \mathscr{A} \rangle$ puede ahora dotarse con una topología apropiada, como sigue. Cualquier n-carta g de S se dice $\mathscr{A}$-compatible si, para todo $f \in \mathscr{A}$, las aplicaciones compuestas $f \circ g^{-1}$ y $g \circ f^{-1}$ son lisas en cada punto donde estén definidas. Es fácil demostrar que, si g y h son cartas $\mathscr{A}$-compatibles de S, tanto $h \circ g^{-1}$ como $g \circ h^{-1}$ son lisas donde quiera estén definidas.[5] Por consiguiente, el conjunto de todas las cartas $\mathscr{A}$-compatibles de S es un n-atlas diferenciable de S, el *atlas máximo* $\mathscr{A}_{\max}$ generado por el atlas dado $\mathscr{A}$. Dotemos a S de la topología más gruesa en virtud de la cual todas las cartas en $\mathscr{A}_{\max}$ son continuas. Con esta topología, S es una variedad topológica n-dimensional.[6]

Sea $U \subseteq S$ un abierto de S; llamemos $\mathscr{A}_{\max}\restriction_U$ a la colección de las restricciones a U de cada carta en $\mathscr{A}_{\max}$. $\mathscr{A}_{\max}\restriction_U$ es entonces el atlas máximo para U y la estructura $\langle U, \mathscr{A}_{\max}\restriction_U \rangle$ es una variedad n-dimensional. Con la estructura de variedad que le confiere $\mathscr{A}_{\max}\restriction_U$, U es una *subvariedad abierta* de S. El lector puede controlar su comprensión de las definiciones precedentes demostrando que, con la estructura recién descrita, el dominio de cualquier n-carta de S es una subvariedad abierta de S. De ahora en adelante, cada vez que mencione una carta de una variedad n-dimensional dada me estaré refiriendo a una n-carta del atlas máximo que la caracteriza como variedad diferenciable.

Sea M una variedad m-dimensional y N una variedad n-dimensional (donde m puede o no ser igual a n). Una aplicación $f: M \to N$ es *lisa en el punto $p \in M$* si hay cartas h y g definidas, respectivamente, en p y en $f(p)$, tales que la aplicación compuesta $g \circ f \circ h^{-1}$ sea lisa en $h(p)$. Se demuestra fácilmente que esta propiedad no depende de la elección de h y g (cf. la nota 5, p. 238). f es *lisa* si es lisa en cada punto de M. Si f es un homeomorfismo liso con inversa lisa f^{-1}, f es un *difeomorfismo*. M y N son variedades *difeomorfas* si hay un difeomorfismo $f: M \to N$. Ello presupone que $m = n$.

7.9 Espacios tangentes

Sea M una variedad n-dimensional. A cada punto $p \in M$ le asociaremos un espacio vectorial real n-dimensional T_pM que llamamos *el espacio tangente* (a M) *en* p.

Considérese la colección $\mathcal{F}(p)$ de todos los campos escalares (esto es, funciones lisas con valores reales) definidos en algún entorno de p. Cualquier función $g \in \mathcal{F}(p)$ puede multiplicarse por un número real α según la regla: $\alpha g(q) = \alpha(g(q))$ para cada $q \in \operatorname{dom} g \subseteq M$. Si g y h pertenecen a $\mathcal{F}(p)$, la suma $g+h$ está definida en la intersección de sus dominios por la condición siguiente: $(g+h)(q) = g(q) + h(q)$ para cada $q \in \operatorname{dom} g \cap \operatorname{dom} h$. Evidentemente, αg y $g+h$ son funciones lisas definidas en un entorno de p y pertenecen, por lo tanto, a $\mathcal{F}(p)$. Provisto de estas operaciones, $\mathcal{F}(p)$ es un espacio vectorial real (de infinitas dimensiones).[7] Como vimos en la p. 162, asociado a $\mathcal{F}(p)$ hay otro espacio vectorial, el espacio dual $\mathcal{F}^*(p)$, que abarca todas las funciones lineales sobre $\mathcal{F}(p)$. Definiremos el espacio tangente T_pM como subespacio de $\mathcal{F}^*(p)$.

Una *curva* en M es una aplicación lisa $\gamma\colon \mathbf{I} \to M$ (donde $\mathbf{I} \subseteq \mathbb{R}$ es un intervalo abierto). El recorrido $\gamma[\mathbf{I}] \subseteq M$ de la aplicación γ es el *camino* de la curva. Diré que esta *pasa por* los puntos que hay en su camino. El argumento variable $u \in \mathbf{I} \subseteq \mathbb{R}$ de la aplicación γ es el *parámetro* de la curva. Adviértase que muchas curvas diferentes pueden pasar por los mismos puntos. Si las curvas $\gamma_1\colon \mathbf{I}_1 \to M$ y $\gamma_2\colon \mathbf{I}_2 \to M$ comparten un mismo camino, hay un homeomorfismo $f : \mathbf{I}_1 \to \mathbf{I}_2$, tal que $\gamma_2 \circ f = \gamma_1$. f es una *reparametrización* de γ_1; en tal caso suele también decirse que cada curva γ_k es una reparametrización de la otra ($k = 1,2$). Consideremos ahora un punto cualquiera p en el camino de la curva $\gamma : \mathbf{I} \to M$. Entonces, $p = \gamma(\alpha)$ para algún $\alpha \in \mathbf{I}$. La *tangente* $\dot{\gamma}_p$ a la curva γ en p es la función lineal definida en $\mathcal{F}(p)$ por

$$\dot{\gamma}_p f = \left.\frac{\mathrm{d}f \circ \gamma}{\mathrm{d}u}\right|_{\alpha} \tag{3}$$

para cada $f \in \mathcal{F}(p)$. Obviamente, $\dot{\gamma}_p \in \mathcal{F}^*(p)$. Para simplificar las

explicaciones, supondré que en el caso de cualquier curva γ mencionada en este apéndice, y para todo punto p en su camino, $\dot{\gamma}_p \neq \mathbf{0}$. Si $p = \gamma(\alpha)$, podemos escribir $\dot{\gamma}(\alpha)$ en vez de $\dot{\gamma}_p$ o $\dot{\gamma}_{\gamma(\alpha)}$.

Considérese el conjunto de todas las curvas que pasan por p. Puede mostrarse que sus respectivas tangentes en p generan un subespacio n-dimensional de $\mathcal{F}^*(p)$. Este subespacio es el *espacio tangente* a M en p, que denotamos con T_pM. Los elementos de T_pM se llaman *vectores tangentes* (a M) en p.

7.10 Tensores, campos tensoriales y sus componentes

Cualquier espacio vectorial $\mathcal{V}$ está asociado automáticamente a otros espacios vectoriales, tales como el espacio dual $\mathcal{V}^*$ formado por las funciones lineales sobre $\mathcal{V}$, y los espacios de las funciones multilineales “covariantes” sobre $\mathcal{V}$, “contravariantes” sobre $\mathcal{V}^*$ y “mixtas” sobre ambos. Esto se aplica naturalmente a cada espacio tangente T_pM de una variedad diferenciable M. En particular, el espacio dual T_pM^* es el *espacio cotangente en* p, y sus elementos se conocen como *covectores en* p. Un *tensor contravariante de tipo* $(q,0)$ en p es una función q-lineal sobre T_pM^*. Un *tensor covariante de tipo* $(0,r)$ en p es una función r-lineal sobre T_pM. Un *tensor mixto de tipo* (q,r) en p es una función $(q+r)$-lineal sobre q copias de T_pM^* y r copias de T_pM. Por ejemplo, una función $(q+r)$-lineal

$$\tau : \underbrace{\mathrm{T}_pM^* \times \ldots \times \mathrm{T}_pM^*}_{q \text{ veces}} \times \underbrace{\mathrm{T}_pM \times \ldots \times \mathrm{T}_pM}_{r \text{ veces}} \to \mathbb{R}$$

es un *tensor mixto de tipo* (q,r) en p. El número $q+r$ es el *rango* del tensor (también si $q = 0$ o $r = 0$). En virtud de estas definiciones, un vector tangente a M en p es un tensor de tipo (1,0) en p, y un covector en p es un tensor de tipo (0,1) en p. Uno puede también referirse a un número real como “un tensor de tipo (0,0)” en p. A menos que q o r sean iguales a 0, nuestra definición de tensor de tipo (q,r) es

ambigua, por cuanto no distingue, por ejemplo, entre funciones trilineales sobre $\mathrm{T}_pM^* \times \mathrm{T}_pM^* \times \mathrm{T}_pM$, $\mathrm{T}_pM \times \mathrm{T}_pM^* \times \mathrm{T}_pM^*$ y $\mathrm{T}_pM^* \times \mathrm{T}_pM \times \mathrm{T}_pM^*$. Para señalar la diferencia entre tres tensores de tipo (2,1) definidos respectivamente sobre estos dominios, se dice que son tensores de rango 3, el primero de los cuales es contravariante en el primer y el segundo índice y covariante en el tercero, mientras que el segundo es covariante en su primer índice y contravariante en el segundo y el tercero, y el tercero es contravariante en el primer y el tercer índice y covariante en el segundo.[8] La extensión de esta terminología al caso general es obvia. Con todo, de ahora en adelante y a menos que del contexto se desprenda otra cosa, entenderemos que un tensor cualquiera de tipo (q,r) es un tensor de rango $(q+r)$ contravariante en sus primeros q índices y covariante en los últimos r índices. Señalo, para ulterior referencia, que un tensor τ de tipo (2,0) o (0,2) se dice *simétrico* si $\tau(v,w) = \tau(w,v)$ para todo par $\langle v,w\rangle \in \mathrm{dom}\ \tau$.

Hay un modo natural de organizar en una variedad k-dimensional (para un entero idóneo k) todos los tensores de un tipo determinado (sin ambigüedad) asociados a todos los puntos de M (vgr. todas las funciones bilineales sobre T_pM para todo $p \in M$). Una estructura $\langle B,M,\pi\rangle$, tal que B y M son variedades diferenciables y π es una aplicación lisa de B en M es un *fibrado* (liso) con base M; π es la *proyección* del fibrado; si $p \in M$, el conjunto $\{x \in B : \pi(x) = p\}$ es la *fibra de* π *sobre* p. Una *sección* del fibrado $\langle B,M,\pi\rangle$ es una aplicación lisa $f: M \to B$, tal que $\pi \circ f = 1_M$, la aplicación idéntica sobre M. Esto significa que a cada argumento $p \in M$ de la sección f, esta asigna como valor precisamente un elemento de la fibra de π sobre p.

Consideremos ahora los tensores de tipo (1,0) en cada punto de la variedad n-dimensional M, esto es, los vectores tangentes a M en cada punto. Pueden organizarse en una variedad $2n$-dimensional de la manera siguiente: Para cada $p \in M$ elegimos una carta x definida en p. Como explico en el párrafo subsiguiente, la carta x determina una base del espacio tangente T_pM, formada por los vectores $\left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p$ allí definidos; cada vector $v \in \mathrm{T}_pM$ es igual a una combinación lineal $v^1 \left.\frac{\partial}{\partial x^1}\right|_p + \dots + v^n \left.\frac{\partial}{\partial x^n}\right|_p$ de estos n vectores. Dejemos que v recorra cada

espacio tangente T_qM, mientras q recorre el dominio de la carta x. La aplicación $v \mapsto \langle x^1(q),\ldots,x^n(q),v^1,\ldots,v^n\rangle$ es obviamente una $2n$-carta del conjunto $TM = \{v : v \in T_pM,\ p \in \text{dom } x\}$, en el sentido explicado en la §7.7. La colección de todas las $2n$-cartas definibles de este modo cuando x recorre el atlas máximo de M constituye un atlas para TM que le confiere la estructura de una variedad $2n$-dimensional. La proyección π: $TM \rightarrow M$ le asigna a cada objeto $v \in TM$ el punto $\pi(v)$ en que v es tangente a M. La estructura $\langle TM,M,\pi\rangle$ es el *fibrado tangente* sobre M. Sea φ una sección de este fibrado; φ le asigna a cada $p \in M$ un vector perteneciente a T_pM, el espacio tangente a M en p; φ es lo que se llama un *campo vectorial* sobre M. Análogamente, una sección ω del fibrado cotangente $\langle TM^*,M,\pi\rangle$ asigna a cada $p \in M$ un covector perteneciente T_pM^*, el espacio cotangente en p; ω es lo que llamamos un *campo de covectores* sobre M. En general, un *campo tensorial* de tipo (q,r) sobre M es una sección del fibrado $\langle B,M,\pi\rangle$, donde B es la variedad diferenciable formada organizando de un modo análogo al descrito todos los tensores de tipo (q,r) en cada punto de M. Un campo vectorial sobre M es, por cierto, lo mismo que un campo tensorial de tipo (1,0) sobre M. Las funciones lisas $M \rightarrow \mathbb{R}$ o *campos escalares* a veces se llaman campos tensoriales de tipo (0,0).

Sea W un campo tensorial de tipo (q,r) sobre M $(q + r > 1)$, y sean V^h y ω^k, respectivamente, un campo vectorial y un campo de covectores sobre M $(1 \leq h \leq q;\ 1 \leq k \leq r)$. W opera sobre el $(q+r)$-tuplo $\langle\omega^1,\ldots,\omega^q,V^1,\ldots,V^r\rangle$ como sigue: W asigna a cada punto p de la variedad M una función $(q+r)$-lineal W_p sobre $(T_pM^*)^q \times (T_pM)^r$; por otra parte, los campos V^h y ω^k asignan, respectivamente, a p las funciones lineales $V^h{}_p$, pertenecientes a $T_pM^{**} = T_pM$, y $\omega^k{}_p$, pertenecientes a T_pM^*. La aplicación $p \mapsto W_p(\omega^1{}_p,\ldots,\omega^q{}_p,V^1{}_p,\ldots,V^r{}_p)$ es un campo escalar sobre M —determinado por los campos W, $\omega^1,\ldots,\omega^q$, $V^1,\ldots$, y V^r— que designo con $W(\omega^1,\ldots,\omega^q,V^1,\ldots,V^r)$.[9] Por analogía con esto, en el caso en que V es un campo tensorial de tipo (1,0), es decir, un campo vectorial ordinario sobre M, y f es un campo escalar, denoto con Vf el campo escalar definido por $p \mapsto V_pf$.

Un campo tensorial sobre M suele ser representado por sus *componentes* relativos a una carta cualquiera. Estos son campos escalares definidos por la operación del campo tensorial sobre ciertos vectores

y covectores determinados por esa carta en cada punto de su dominio. Para explicarlo, debo llamar la atención del lector sobre unos cuantos hechos obvios. Sea x una carta definida en un abierto U_x de una variedad n-dimensional M. Sea x^k la k-ésima coordenada de la carta x $(1 \leq k \leq n)$. Cada punto $p \in U_x$ está sobre el camino de una curva γ tal que (i) γ está definida en un intervalo abierto incluido en el recorrido de x^k; (ii) γ aplica este intervalo inyectivamente en U_x; (iii) para cada punto q en el camino de γ, $q = \gamma \circ x^k(q)$, y (iv) para cada punto q en el camino de γ y para cada índice $h \neq k$, $x^h(q) = x^h(p)$ $(1 \leq h \leq n)$. Hay siempre varias curvas que pasan por p y cumplen estas cuatro condiciones, pero una de ellas es máxima, en cuanto su dominio incluye el de las otras. La curva máxima que pasa por p y cumple las condiciones (i)–(iv) es la *curva paramétrica* de la coordenada x^k por p, o *k-ésima curva paramétrica* de la carta x por el punto p. La tangente en p a la curva paramétrica de x^k por p se designa con el símbolo $\left.\frac{\partial}{\partial x^k}\right|_p$. Los n vectores tangentes $\left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p$ son linealmente independientes y por tanto constituyen una base del espacio tangente T_pM. Las curvas paramétricas por p de la carta x nos dan asimismo una base del espacio cotangente T_pM^*. En efecto, si $dx^k{}_p$ designa el covector en p definido por

$$dx^k{}_p(v) = vx^k \tag{4}$$

para cada $v \in T_pM$, entonces los covectores $dx^1{}_p, \ldots, dx^k{}_p$ constituyen precisamente la base dual de $\left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p$, pues

$$dx^k_p\left(\left.\frac{\partial}{\partial x^h}\right|_p\right) = \left.\frac{\partial x^k}{\partial x^h}\right|_p = \delta^k{}_h = \begin{cases} 1 & \text{si } k = h \\ 0 & \text{si } k \neq h \end{cases} \tag{5}$$

Ahora puedo explicar qué son los *componentes* de un campo tensorial relativos a una carta dada. Considérese, por ejemplo, un campo tensorial R de tipo (1,3) sobre la variedad n-dimensional M. Sea x una carta definida en el abierto $U_x \subseteq M$. Para cada cuádruplo $\langle i,j,k,h \rangle$

$(1 \leq i,j,k,h \leq n)$ definimos un campo escalar $R^i{}_{jkh}$ sobre U_x por

$$R^i{}_{jkh}: p \mapsto R_p\left(\mathrm{d}x^i_p, \left.\frac{\partial}{\partial x^j}\right|_p, \left.\frac{\partial}{\partial x^k}\right|_p, \left.\frac{\partial}{\partial x^h}\right|_p\right) \tag{6}$$

$R^i{}_{jkh}$ es el (i,j,k,h)-ésimo componente de R relativo a la carta x. Como sabemos, R es una sección del fibrado de todos los tensores de tipo (1,3) en cada punto de M. En otras palabras, R asigna a cada $p \in M$ una función cuadrilineal R_p sobre $\mathrm{T}_pM^* \times \mathrm{T}_pM \times \mathrm{T}_pM \times \mathrm{T}_pM$. Como cada elemento de T_pM es igual a una combinación lineal de los n vectores tangentes, $\left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p$, y cada elemento de T_pM^* es igual a una combinación lineal de los n vectores cotangentes $\mathrm{d}x^1{}_p, \ldots,$ $\mathrm{d}x^k{}_p$, es claro que la función cuadrilineal R_p está completamente determinada en cada punto $p \in U_x$ por los n^4 componentes definidos por las ecuaciones (6). Conforme a la convención adoptada en la p. 170 (después de la llamada 9), el i-ésimo componente V^i de un campo vectorial V es el campo escalar $p \mapsto V_p x^i$.

Sea f: $M \to N$ una aplicación lisa de una variedad m-dimensional en una variedad n-dimensional. f determina en cada $p \in M$ una aplicación lineal $f_{*p}: \mathrm{T}_pM \to \mathrm{T}_{f(p)}N$. Esta aplicación está completamente determinada por los valores que asigna a los vectores de una base de T_pM. Si x e y son, respectivamente, una carta de M definida en p y una carta de N definida en $f(p)$, f_{*p} queda definida por:

$$f_{*p}\left(\left.\frac{\partial}{\partial x^k}\right|_p\right) = \left.\frac{\partial(y^h \circ f)}{\partial x^k}\right|_p \left.\frac{\partial}{\partial y^h}\right|_{f(p)} \tag{7}$$

Como f_{*p} es una aplicación lineal de un espacio vectorial en otro, da lugar a una aplicación "retroactiva" del dual del segundo espacio en el dual del primero. Esta aplicación, habitualmente designada con $f^*_p: \mathrm{T}_{f(p)}N^* \to \mathrm{T}_pM^*$.[10] queda definida como sigue en términos de la cartas x e y:

$$f_p^*\left(\mathrm{d}y^k_{f(p)}\right) = \left.\frac{\partial(y^k \circ f)}{\partial x^h}\right|_p \mathrm{d}x^h_p \tag{8}$$

La aplicación $f{:}M \to N$ determina asimismo aplicaciones lineales de cada espacio de tensores contravariantes asociado a T_pM al espacio de tensores del mismo tipo asociado a $\mathrm{T}_{f(p)}N$, y de cada espacio de tensores covariantes asociado a $\mathrm{T}_{f(p)}N$ al espacio correspondiente asociado a T_pM. Estas aplicaciones se designan con f^*_p y f_{*p}, respectivamente, y se definen, *mutatis mutandis*, por analogía con (7) y (8). Por ejemplo, si ω es un campo tensorial de tipo (0,3) sobre N y X, Y y Z son campos vectoriales sobre M,

$$f_p^*\left(\omega_{f(p)}\right)\left(X_p, Y_p, Z_p\right) = \omega_{f(p)}\left(f_{*p}(X_p), f_{*p}(Y_p), f_{*p}(Z_p)\right) \tag{9}$$

donde $f_p^*\left(\omega_{f(p)}\right)$ es el valor que f_p^* le asigna a $\omega_{f(p)}$ y por tanto un tensor de tipo (0,3) en p.

7.11 Producto tensorial y contracción

En esta sección explicaré brevemente dos operaciones algebraicas sobre campos tensoriales. Como el tema es seco como yesca, recomiendo al lector que se salte esta sección y luego retorne a ella cuando sienta que lo necesita. También es aconsejable, antes de consultarla, leer en la nota 9, pp. 238–239, la definición de $\mathcal{V}^q{}_r(M)$, el módulo —sobre el anillo $\mathcal{F}(M)$— de los campos tensoriales de tipo (q,r) sobre M.

Si $X \in \mathcal{V}^q{}_r(M)$ e $Y \in \mathcal{V}^t{}_s(M)$, el *producto tensorial* $X \otimes Y$ es el campo tensorial de tipo $(q+t,r+s)$ —contravariante en los índices del primero al q-ésimo y del $(q+r+1)$-ésimo al $(q+r+t)$-ésimo, y covariante en los índices $(q+1)$-ésimo al $(q+r)$-ésimo y $(q+r+t+1)$-ésimo al $(q+r+t+s)$-ésimo— que le asigna a cada $p \in M$ la función $(q+r+t+s)$-lineal $(X \otimes Y)_p$, definida por

$$(X \otimes Y)_p(\omega_1,\ldots,\omega_q,V_1,\ldots,V_r,\psi_1,\ldots,\psi_t,W_1,\ldots,W_s) = X_p(\omega_1,\ldots,\omega_q,V_1,\ldots,V_r)Y_p(\psi_1,\ldots,\psi_t,W_1,\ldots,W_s) \qquad (10)$$

donde $V_1,\ldots,V_r,W_1,\ldots,W_s$ son cualesquiera vectores de T_pM y $\omega_1,\ldots,\omega_q$, $\psi_1,\ldots,\psi_t$ son cualesquiera covectores de T_pM^*. Evidentemente, esta operación es asociativa: $(X \otimes Y) \otimes Z = X \otimes (Y \otimes Z)$. Los componentes del producto tensorial se calculan fácilmente a partir de los componentes de sus factores. Por ejemplo, si T y ω son campos tensoriales de tipo (1,1) y (0,2), respectivamente, y $T \otimes \omega$ es el campo tensorial R de tipo (1,3) cuyos componentes definimos en las ecuaciones (6). Entonces, el (i,j,k,h)-ésimo componente R^i_{jkh} de R relativo a un carta dada x está dado por

$$\begin{aligned} R^i_{jkh}(p) &= R_p\left(\mathrm{d}x^i_p, \left.\frac{\partial}{\partial x^j}\right|_p, \left.\frac{\partial}{\partial x^k}\right|_p, \left.\frac{\partial}{\partial x^h}\right|_p\right) = \\ &= T_p\left(\mathrm{d}x^i_p, \left.\frac{\partial}{\partial x^j}\right|_p\right)\omega_p\left(\left.\frac{\partial}{\partial x^k}\right|_p, \left.\frac{\partial}{\partial x^h}\right|_p\right) = T^i_j(p)\omega_{kh}(p) \end{aligned} \qquad (11)$$

Obviamente, algunos campos tensoriales de tipo (q,r) sobre M son iguales al producto tensorial de q campos vectoriales y r campos de covectores sobre M. Llamémosles campos tensoriales *simples* de tipo (q,r). Todo los demás campos tensoriales de tipo (q,r) sobre M pueden expresarse como combinaciones lineales de los campos tensoriales simples. Por tanto, dichos campos simples generan el módulo $\mathcal{V}^q_r(M)$. Obsérvese, en particular, que si $U_x \subseteq M$ es el dominio de una carta x, el módulo $\mathcal{V}^q_r(U_x)$ es generado por los productos tensoriales

$$\partial/\partial x^{i_1} \otimes \ldots \otimes \partial/\partial x^{i_q} \otimes \mathrm{d}x^{k_1} \otimes \ldots \otimes \mathrm{d}x^{k_r}$$
$$(1 \leq i_j,k_h \leq n;\ 1 \leq j \leq q;\ 1 \leq h \leq r)$$

Por tanto, la restricción a U_x de, digamos, el campo tensorial R de tipo (1,3) de las ecuaciones (6) es igual a una combinación lineal de los productos tensoriales $\partial/\partial x^i \otimes \mathrm{d}x^j \otimes \mathrm{d}x^k \otimes \mathrm{d}x^h$ $(1 \leq i,j,k,h \leq n)$. No es difícil demostrar que, en dicha combinación lineal, el coeficiente escalar de cada término $\partial/\partial x^i \otimes \mathrm{d}x^j \otimes \mathrm{d}x^k \otimes \mathrm{d}x^h$ no es otro

que el (i,j,k,h)-ésimo componente de R relativo a x, esto es, el campo escalar R^i_{jkh} definido en la ecuación (6). En otras palabras,

$$R\restriction U_x = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n} R^i_{jkh}\frac{\partial}{\partial x^i}\otimes \mathrm{d}x^j\otimes \mathrm{d}x^k\otimes \mathrm{d}x^h \tag{12}$$

En la vida real, claro, físicos y matemáticos escriben simplemente R en vez de $R\restriction U_x$, porque el lado derecho de la ecuación no da lugar a dudas sobre el dominio considerado, y simplifican también drásticamente el lado derecho gracias a la *convención de Einstein* (p. 107), en virtud de la cual la sumatoria se sobreentiende sobre cada índice que aparezca dos veces en un mismo término. La ecuación (12) entonces se formula corrientemente así:

$$R = R^i_{jkh}\frac{\partial}{\partial x^i}\otimes \mathrm{d}x^j\otimes \mathrm{d}x^k\otimes \mathrm{d}x^h \tag{12′}$$

Para comprobar que domina esta notación, el lector demuestre la ecuación (12), desarrollándola para el caso $n = 2$.

La *contracción* de un campo tensorial de tipo (q,r) sobre M respecto a sus índices k-ésimo y h-ésimo ($1 \leq k \leq q$; $q+1 \leq h \leq q+r$) es la aplicación lineal $\mathrm{C}^k_h{:}\mathcal{V}^q_r(M) \to \mathcal{V}^{q-1}_{r-1}(M)$ cuyo valor sobre un campo tensorial simple cualquiera sobre M de tipo (q,r), $W = V_1 \otimes \ldots \otimes V_r \otimes \omega_1 \otimes \ldots \otimes \omega_q$, es el campo tensorial de tipo $(q-1,r-1)$:

$$\begin{aligned}\mathrm{C}^k_h(W) &= \mathrm{C}^k_h(V_1\otimes\ldots\otimes V_q\otimes\omega_1\otimes\ldots\otimes\omega_r) = \\ &= \omega_h\left(V_k\right)(V_1\otimes\ldots\otimes V_{k-1}\otimes V_{k+1}\otimes\ldots\otimes V_q\otimes \\ &\qquad \omega_1\otimes\ldots\otimes\omega_{h-1}\otimes\omega_{h+1}\otimes\ldots\otimes\omega_r)\end{aligned} \tag{13}$$

La generalización de (13) al caso de una combinación lineal cualquiera de campos tensoriales simples de tipo (q,r) no presenta dificultades. Los componentes de $\mathrm{C}^k_h(W)$ relativos a una carta dada pueden fácilmente calcularse a partir de los componentes de W. Consideremos una vez más el campo tensorial R de tipo (1,3) de la ecuación (12). La contracción $\mathrm{C}^1_4(R)$ de R respecto a sus índices primero y cuarto es un campo tensorial de tipo (0,2) que llamaré $\widehat{R}$. En el dominio U_x

de nuestra carta x, R es igual a la combinación lineal que figura al lado derecho de la ecuación (12). Por lo tanto, por (13) y (5),

$$\begin{aligned}\widehat{R} = \mathrm{C}^1_4(R) &= R^i_{jkh}\mathrm{d}x^h\left(\frac{\partial}{\partial x^i}\right)(\mathrm{d}x^j \otimes \mathrm{d}x^k) = \\ &= \delta^h_i R^i_{jkh}(\mathrm{d}x^j \otimes \mathrm{d}x^k) = R^h_{jkh}(\mathrm{d}x^j \otimes \mathrm{d}x^k)\end{aligned} \tag{14}$$

donde uso la convención de Einstein, como siempre haré hasta el final de este apéndice. Por lo tanto, los componentes de $\widehat{R}$ relativos a la carta x son los campos escalares

$$\widehat{R}_{jk} = R^h_{jkh} \tag{15}$$

7.12 Transporte paralelo, geodésicas y curvatura

Sea M una variedad n-dimensional. En la §7.12 y la §7.13, a menos que indique otra cosa, cuando hablo de un campo tensorial, vectorial o escalar me refiero a uno definido en una subvariedad abierta de M. Bosquejaré modos de enriquecer la variedad diferenciable M con más estructura. En la presente sección, muestro como es posible privilegiar una clase de curvas en M, que, por razones históricas, se llaman *geodésicas*;[11] también defino el concepto de *curvatura*, que depende del mismo concepto que utilizaré para caracterizar las geodésicas. En la §7.13 me ocupo con las métricas riemannianas. Las geodésicas se concibieron inicialmente desde el punto de vista métrico, como curvas de longitud extrema entre dos puntos; pero Levi-Civita (1917) mostró que dependen de una correspondencia ("paralelismo") entre vectores tangentes a M en distintos puntos, cuya definición no presupone una métrica.

Como todos los espacios tangentes a M son espacios vectoriales reales n-dimensionales y por tanto isomorfos entre sí y con el espacio vectorial $\mathbb{R}^n$, todo lo que necesitamos para establecer una correspondencia entre vectores tangentes a M en distintos puntos es definir un

isomorfismo entre T_pM y T_qM para cada par de puntos p, $q \in M$. La forma obvia de hacerlo es elegir una base en cada punto, convenir un orden entre sus elementos y poner las bases ordenadas en correspondencia mutua. Una base ordenada del espacio tangente T_pM es una n-ada en p. Sea FM el conjunto de todas las n-adas en todos los puntos de M. La proyección $\pi : FM \to M$ asigna a cada n-ada en FM el punto de M de cuyo espacio tangente esa n-ada es una base. FM deviene una variedad $(n+n^2)$-dimensional si, para cada carta x of M, con dominio U_x, definimos en $\pi^{-1}[U_x]$ una carta correspondiente X con coordenadas $X^1,\ldots,X^{n^2+n}$ que satisfacen la condición siguiente: si $F = \langle F_1,\ldots,F_n\rangle$ es una n-ada en $p \in U_x$, entonces

$$X^i(F) = x^i(p) \qquad X^{nk+i}(F) = F_k(x^i) \qquad (1 \leq i,k \leq n) \tag{16}$$

La colección de todas estas cartas X es un atlas para FM. Sea $\mathscr{A}$ el atlas máximo compatible con FM. Claramente, π es lisa en $\langle FM,\mathscr{A}\rangle$. Ahora definiré una acción del grupo lineal general $GL(n,\mathbb{R})$ sobre FM.[12] Sea $\mathsf{B} = (\mathsf{B}^i_j)$ una matriz real no singular $n \times n$ y sea $F = \langle F_1,\ldots,F_n\rangle$ una n-ada en $p \in M$; entonces $\mathsf{B}F = \langle \mathsf{B}^k{}_1 F_k,\ldots,\mathsf{B}^k{}_n F_k\rangle$. $\mathsf{B}F$ es obviamente una n-ada en p. Nótese que $\mathsf{B}F$ pertenece a la misma fibra de π que F; claramente, si F y F' son dos n-adas en el mismo punto siempre hay una matriz A del género descrito tal que $F' = \mathsf{B}F$; más aún, $\mathsf{B}F = F$ si y solo si $\mathsf{B} = \mathrm{diag}(1,\ldots,1)$, esto es, si para los n índices k y h, $\mathsf{B}^k{}_h = \delta^k{}_h$. La estructura $\langle FM,M,\pi,GL(n,\mathbb{R})\rangle$ definida por las estipulaciones precedentes se llama el *fibrado principal de n-adas sobre M*. Lo designaré con FMG. Una sección de FMG es un campo de n-adas sobre M (Élie Cartan lo llamó *un repère mobile*). Si tal campo de n-adas existe, determina isomorfismos entre todos los pares de espacios tangentes M. Dos vectores tangentes a M en distintos puntos son imagen el uno del otro por el isomorfismo pertinente si son ambos la misma combinación lineal del valor local del campo de n-adas en los puntos respectivos. Si esta condición se cumple, se dice que los dos vectores son *paralelos*. La variedad M es *paralelizable* si y solo si existe una sección del fibrado principal FMG.

Generalmente, las variedades n-dimensionales no son paraleliza-

bles.[13] No obstante, Levi-Civita pudo definir una noción de *paralelismo a lo largo de una curva dada* que es aplicable en cualquier variedad n-dimensional. Su idea ha sido articulada de varios modos. Creo que la más natural es la que propuso Ehresmann (1950). Pero ella envuelve tecnicismos que no cabría explicar aquí.[14] Opto pues por la formulación de Koszul, publicada por Nomizu (1954); es más artificial, pero es más fácil explicarla y trabajar con ella.

Una *conexión lineal* ∇ (de Koszul) sobre M asigna a cada par de $\langle X,Y\rangle$ de campos vectoriales un campo vectorial $\nabla_X Y$ sobre dom $X \cap$ dom Y, que satisface las siguientes ecuaciones, para cualesquiera campos vectoriales X, Y, Z y W y cada campo escalar f, dondequiera que estén definidas las expresiones que figuran a cada lado:

$$\begin{aligned} \nabla_X(Z + W) &= \nabla_X Z + \nabla_X W \\ \nabla_{X+Y} Z &= \nabla_X Z + \nabla_Y Z \\ \nabla_{fX} Z &= f\nabla_X Z \\ \nabla_X fZ &= (Xf)Z + f\nabla_X Z \end{aligned} \tag{17}$$

$\nabla_X Y$ es la *derivada covariante de Y en la dirección de X*. Considérese una carta x definida sobre $U_x \subseteq M$. Para facilitar la escritura, designo con $\nabla_k Y$ a la derivada covariante de Y en la dirección de $\partial/\partial x^k$. $\nabla_X Y$ puede expresarse en $U_x \cap$ dom $X \cap$ dom Y como una combinación lineal de los $\partial/\partial x^k$ $(1 \leq k \leq n)$. En particular,

$$\nabla_k \frac{\partial}{\partial x^h} = \Gamma^i_{kh} \frac{\partial}{\partial x^i} \tag{18}$$

Los campos escalares Γ^i_{kh} son los *componentes de la conexión lineal ∇ relativos a la carta x*. Como es obvio, sobre U_x la conexión lineal ∇ está completamente determinada por ellos.

Una conexión lineal ∇ sobre M le asigna unívocamente a cada campo vectorial V y cada campo tensorial W de tipo (q,r) $(q,r \geq 0)$ un campo tensorial $\nabla_V W$, de tipo (q,r), llamado la *derivada covariante de W en la dirección de V*, que cumple con las condiciones siguientes, dondequiera que las aplicaciones nombradas estén definidas:

(i) Si α, $\beta \in \mathbb{R}$ y U y W son campos tensoriales de tipo (q,r),

$\nabla_V(\alpha W+\beta U) = \alpha\nabla_V W + \beta\nabla_V U$.

(ii) Si $\alpha, \beta \in \mathbb{R}$ y Z es un campo vectorial, $\nabla_{\alpha V+\beta Z}W = \alpha\nabla_V W + \beta\nabla_Z W$.

(iii) Si $q = 1$ y $r = 0$, esto es, si W es un campo vectorial, $\nabla_V W$ es precisamente el campo vectorial designado de este modo según nuestra anterior estipulación.

(iv) Si $q = r = 0$, esto es, si W es un campo escalar f, $\nabla_V W = \nabla_V f = Vf$;[15]

(v) Si Z es un campo tensorial de tipo (s,t),

$$\nabla_V(W \otimes Z)= \nabla_V W \otimes Z + W \otimes \nabla_V Z,$$

esto es, un campo tensorial de tipo $(q+s,r+t)$, contravariante en los primeros q índices y en los índices $(q+s+1)$-ésimo al $(q+s+r)$-ésimo, y contravariante en los índices $(q+1)$-ésimo al $(q+s)$-ésimo y en los últimos t.

(vi) ∇_V conmuta con cualquier operación de contracción C:

$$\nabla_V \circ \mathrm{C} = \mathrm{C} \circ \nabla_V$$

Si W es un campo tensorial de tipo (q,r), la *derivada covariante* (absoluta) de W —simbolizada ∇W— es el campo tensorial de tipo $(q,r+1)$ definido sobre dom W por

$$\nabla W(\omega_1,\ldots,\omega_q,V_1,\ldots,V_r,V) = (\nabla_V W)(\omega_1,\ldots,\omega_q,V_1,\ldots,V_r) \qquad (19)$$

donde V, $V_1,\ldots,V_r$ son cualesquiera campos vectoriales y $\omega_1,\ldots,\omega_q$ son cualesquiera campos de covectores sobre dom W.

Si ∇ es una conexión lineal sobre una variedad n-dimensional M podemos usarla para definir una relación de paralelismo entre vectores a lo largo de cualquier curva en M. Sea $\gamma: (a,b) \to M$ una curva definida en el intervalo abierto $(a,b) \subseteq \mathbb{R}$. Llamo campo vectorial sobre γ a un campo vectorial definido sobre un entorno abierto de cada punto en el camino de γ. Hay incontables campos vectoriales diferentes sobre γ cuyo valor en cada punto p del camino de γ es igual a la tangente $\dot{\gamma}_p$ de γ en ese punto (definida en la p. 167, ecuación (3)). Cualquiera de ellos sirve para nuestro propósito. Denotémosle con $\dot{\gamma}$. Un campo vectorial V se dice *paralelo a lo largo de la curva* γ si la derivada covariante $\nabla_{\dot{\gamma}}V$ of V en la dirección de $\dot{\gamma}$ es igual a 0 en cada punto del camino de γ donde V esté definido. La curva γ es una *geodésica*

si y solo si el campo vectorial $\dot{\gamma}$ es paralelo a lo largo de γ.[16]

Enunciaré ahora estas definiciones en términos de componentes. Sea x una carta definida en un abierto U_x que contiene todo el recorrido de la curva γ: $(a,b) \to M$. Sea V un campo vectorial sobre γ. Escribiré V^i por Vx^i. El campo escalar V^i, definido sobre $U_x \cap \text{dom } V$, es el i-ésimo componente de V relativo a la carta x (cf. p. 172). Las funciones compuestas $v^k = V^k \circ \gamma$ $(1 \leq k \leq n)$ son funciones reales lisas cuyo dominio es el intervalo (a,b). V es paralelo a lo largo de γ si y solo si la funciones v^k son soluciones del sistema

$$\frac{dv^k}{dt} + v^i \frac{d(x^j \circ \gamma)}{dt}\left(\Gamma^k_{ji} \circ \gamma\right) = 0 \tag{20}$$

donde t es el parámetro de la curva γ y las Γ^k_{ji} son los componentes de ∇ relativos a la carta x según fueron definidos en las ecuaciones (18). A su vez, esto implica que γ es una geodésica si y solo si

$$\frac{d^2(x^k \circ \gamma)}{dt^2} + \frac{d(x^i \circ \gamma)}{dt}\frac{d(x^j \circ \gamma)}{dt}\left(\Gamma^k_{ji} \circ \gamma\right) = 0 \tag{21}$$

Desde la perspectiva de Ehresmann la estrecha relación entre geodésicas y curvatura puede presentarse en forma muy clara y natural. Desde el punto de vista de Koszul que he adoptado aquí el concepto de curvatura se introduce algo dogmáticamente y para ver su vínculo con las geodésicas hay que atender a la ecuación de la desviación geodésica que enunciaré al final, sin demostrarla. Sean ∇ y M como antes. El *producto de Lie* $[X,Y]$ de los campos vectoriales X e Y es el campo vectorial definido sobre dom $X \cap$ dom Y por la condición

$$[X,Y]f = X(Yf) - Y(Xf) \tag{22}$$

para cualquier campo escalar f: (dom $X \cap$ dom Y) $\to \mathbb{R}$. A cada par $\langle X,Y \rangle$ de campos vectoriales le asignamos un operador lineal $\mathcal{R}(X,Y)$ tal que, para cualquier campo vectorial Z,

$$\mathcal{R}(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z \tag{23}$$

en dom $X \cap$ dom $Y \cap$ dom Z.[17] La *curvatura* R de la variedad M

provista de la conexión lineal ∇ es el campo tensorial de tipo (1,3) que, para cualesquiera campos vectoriales X, Y y Z y campo de covectores ω, satisface la ecuación siguiente

$$R(\omega,X,Y,Z) = \omega(\mathcal{R}(X,Y)Z) \tag{24}$$

en dom $\omega \cap$ dom $X \cap$ dom $Y \cap$ dom Z. R se llama también *el tensor de Riemann*.[18] Si R es idéntico a 0, la conexión ∇ y la variedad M provista de esa conexión se dicen *planas*. Los componentes de R relativos a la carta x de M son precisamente los campos escalares $R^i{}_{jkh}$ definidos en las ecuaciones (6). Teniendo presentes las ecuaciones (5), es posible derivar de (12′) y (19) la siguiente fórmula para los componentes de la curvatura:

$$R^h{}_{ijk} = \frac{\partial \Gamma^h_{ik}}{\partial x^j} - \frac{\partial \Gamma^h_{ij}}{\partial x^k} + \Gamma^s_{ik}\Gamma^h_{sj} - \Gamma^s_{ij}\Gamma^h_{sk} \tag{25}$$

Einstein y sus contemporáneos llaman *tensor de Ricci* a la contracción $C^1{}_4(R)$ de R en sus índices primero y cuarto.[19] Por lo tanto, si en las ecuaciones (14) entendemos que R denota el tensor de Riemann, el tensor de Ricci es el campo tensorial de tipo (0,2) allí denotado por $\widehat{R}$. Pero normalmente no se usan dos símbolos diferentes para el tensor de Riemann y el de Ricci pues no hay peligro de confundirlos si uno los nombra por sus componentes genéricos, $R^h{}_{ijk}$ y R_{ij}, como es habitual.

Paso por alto el concepto de *torsión* que no juega ningún papel en la teoría general de la relatividad. Baste decir que la variedad M, con conexión lineal ∇, se dice *sin torsión* si

$$\nabla_X Y - \nabla_Y X = [X,Y] \tag{26}$$

para cualesquiera campos vectoriales X e Y. Esto implica que los componentes Γ^k_{ij} de ∇ relativos a cualquier carta de M son simétricos en los dos índices inferiores ($\Gamma^k_{ij} = \Gamma^k_{ji}$, siempre). De este hecho y las ecuaciones (21) se infiere que una conexión lineal sin torsión está completamente determinada por sus geodésicas.

Enunciaré sin prueba una importante relación entre curvatura y

geodésicas que vale en una variedad M con conexión lineal ∇ sin torsión. Una *congruencia de curvas* $\mathcal{C}$ en una variedad M es una colección de curvas en M tal que cada punto de M se halla en el camino de una y solo una curva de $\mathcal{C}$. La congruencia $\mathcal{C}$ está asociada a un campo vectorial V cuyo valor V_p en cada $p \in M$ no es otro que el vector $\dot{\gamma}_p \in \mathrm{T}_pM$ tangente en p a la única $\gamma \in \mathcal{C}$ cuyo camino contiene a p. Si $\mathcal{C}$ es una congruencia de geodésicas, entonces su campo vectorial asociado V satisface la ecuación $\nabla_V V = 0$. Si ∇ es plana, V_p es paralelo a V_q, para todo p y $q \in M$, a lo largo de cualquier curva que una p y q. Pero si ∇ no es plana, dicha relación de paralelismo está asegurada solo a lo largo de cada geodésica γ de $\mathcal{C}$. Sabemos que las tangentes de un geodésica $\gamma^1 \in \mathcal{C}$ son todas paralelas entre sí a lo largo de γ^1. Pero cuando transportamos paralelamente una de ellas, digamos $\dot{\gamma}^1_p$, a lo largo de una geodésica θ cuyo camino corta el de γ^1 en p (de modo que $\theta \notin \mathcal{C}$), hasta un punto q en el recorrido de otra geodésica $\gamma^2 \in \mathcal{C}$, nos encontramos con lo siguiente: el vector tangente a M en q que es paralelo a $\dot{\gamma}^1_p$ a lo largo de θ normalmente no coincide con $\dot{\gamma}^2_q$. Se dice, por eso, que aunque cada geodésica $\gamma \in \mathcal{C}$ mantiene característicamente una dirección constante a lo largo de sí misma, esa dirección continuamente converge hacia o diverge de las direcciones sostenidas por las geodésicas vecinas pertenecientes a la congruencia $\mathcal{C}$. Para expresar el grado de divergencia o convergencia y en cierto modo medirlo, introducimos un campo X de "vectores conectores" tal que $[V,X] = 0$ y tomamos la segunda derivada covariante $\nabla_V\nabla_V X$ de X en la dirección de V. Se puede demostrar que

$$\nabla_V\nabla_V X = \nabla_V\nabla_X V = [\nabla_V,\nabla_X]\mathrm{V} + \nabla_X\nabla_V V \tag{27}$$

Como $0 = \nabla_V V = \nabla_X\nabla_V V$ y también $0 = [V,X] = \nabla_{[V,X]}$, la ecuación (23) implica que

$$\nabla_V\nabla_V X = \mathcal{R}(V,X)V \tag{28}$$

La ecuación (28) se llama la *ecuación de desviación geodésica.*[20]

7.13 *Métrica riemanniana*

Un campo tensorial **g** de tipo (0,2) sobre una variedad n-dimensional M es una *métrica riemanniana* sobre M, y la estructura $\langle M,\mathbf{g}\rangle$ es una *variedad riemanniana* (o *espacio riemanniano*) si **g** tiene las siguientes propiedades:

M1 **g** es *simétrica*, esto es, si V y W son dos campos vectoriales cualesquiera, $\mathbf{g}(V,W) = \mathbf{g}(W,V)$ en dom $V \cap$ dom W.

M2 **g** es *no degenerada*, esto es, $\mathbf{g}_p(V,W) = 0$ para cada campo vectorial W y todo $p \in$ dom $V \cap$ dom W si y solo V es idéntico a 0; esto es, si y solo si el campo vectorial V asigna a cada $p \in$ dom V el vector cero de T_pM.

Digo que **g** es una *métrica riemanniana propiamente tal* y que $\langle M,\mathbf{g}\rangle$ es una *variedad riemanniana* (o un *espacio riemanniano*) propiamente tal si, además de **M1** y **M2**, **g** tiene la propiedad siguiente:

M3 **g** es *positiva definida*, esto es, para cada campo vectorial V, $\mathbf{g}_p(V,V) > 0$ para todo $p \in$ dom V, a menos que V sea idéntico a 0, en cuyo caso, el campo escalar $\mathbf{g}(V,V)$ es idéntico a 0.

Para indicar que **g** cumple las condiciones **M1** y **M2**, pero no cumple la condición **M3**, digo a veces que **g** es una métrica riemanniana *impropia*.[21]

Como sabemos, un campo tensorial **g** de tipo (0,2) sobre M asigna a cada $p \in M$ una función bilineal $\mathbf{g}_p$ sobre el espacio tangente T_pM. De la definición precedente se desprende que si **g** es una métrica riemanniana, $\mathbf{g}_p$ es precisamente un producto escalar sobre T_pM; si **g** satisface además la condición **M3**, esto es, si $\langle M,\mathbf{g}\rangle$ es una *variedad riemanniana* propiamente tal, la función bilineal $\mathbf{g}_p$ es por cierto un producto interno (cf. p. 162).

Un producto interno sobre un espacio vectorial proporciona un modo natural de medir la *longitud* de los vectores y el *ángulo* formado por dos cualesquiera de ellos. Si $\langle M,\mathbf{g}\rangle$ es una variedad riemanniana propiamente tal y v y w son vectores tangentes a M en el mismo punto

p, la longitud $\|v\|$ de v y el ángulo $\measuredangle(v,w)$ que forma con w se definen así:

$$\|v\| = \sqrt{\mathbf{g}_p(v,v)} \tag{29}$$

$$\cos \measuredangle(v,w) = \frac{\mathbf{g}_p(v,w)}{\|v\|\|w\|} \tag{30}$$

Evidentemente estas definiciones no sirven si la métrica riemanniana **g** no cumple la condición **M3**, pues entonces la longitud $\|v\|$ de un vector $v \neq 0$ podría ser un número imaginario o igual a 0. Sin embargo, aun en ese caso tiene sentido decir que dos vectores v, $w \in \mathrm{T}_pM$ son *ortogonales* si $\mathbf{g}_p(v,w) = 0$.[22]

Si $\langle M,\mathbf{g}\rangle$ es una variedad riemanniana en mi sentido de este término, pero **g** no cumple la condición **M3**, todos los vectores tangentes a M en cualquier punto p pertenecen a una y solo una de tres clases. Un vector $v \in \mathrm{T}_pM$ se dice

(*a*) *temporaloide*, si $\mathbf{g}_p(v,v) > 0$;
(*b*) *espacialoide*, si $\mathbf{g}_p(v,v) < 0$;
(*c*) *nulo* o *luminoide* si $\mathbf{g}_p(v,v) = 0$.[23]

Con la función bilineal $\mathbf{g}_p$ es posible dar una definición de longitud útil para comparar vectores de la misma clase, siempre que no sean nulos. Simplemente ponemos

$$\|v\| = \sqrt{|\mathbf{g}_p(v,v)|} \tag{31}$$

La clasificación de los vectores permite también clasificar las curvas en $\langle M,\mathbf{g}\rangle$, aunque no exhaustivamente. Una curva $\gamma\colon \mathbf{I} \to M$, se dice *temporaloide*, *espacialoide* o *nula* si en cada punto p de su camino la tangente $\dot{\gamma}_p$ es, respectivamente, un vector temporaloide, espacialoide o nulo. Una curva que no pertenezca a ninguna de estas tres clases se llama *mixta*, pues algunos de sus vectores son de una clase y otros de otra.[24] Si γ no es mixta, la longitud de γ entre los puntos $\gamma(a)$ y $\gamma(b)$ (a, $b \in \mathbf{I}$; $a < b$) esta dada por la integral definida

$$\int_a^b \|\dot{\gamma}(u)\|\mathrm{d}u \tag{32}$$

(donde he escrito $\dot{\gamma}(u)$ por $\dot{\gamma}_{\gamma(u)}$, la tangente a γ en un punto cualquiera $\gamma(u)$ de su camino; cf. p. 168). Es fácil probar que la reparametrización de una curva no afecta su longitud entre dos puntos dados. Si γ es temporaloide o espacialoide es posible reparametrizar γ de modo que $\|\dot{\gamma}(u)\| = 1$ para cada valor de u; entonces

$$\int_a^{a+u} \|\dot{\gamma}(u)\| \mathrm{d}u = \int_a^{a+u} \mathrm{d}u \tag{33}$$

En este caso, pues, el incremento del parámetro u entre dos de sus valores es igual a la longitud del arco descrito por γ mientras u varía de uno al otro, y se dice que γ está parametrizada por la longitud de arco. Una métrica riemanniana no nos da una noción de longitud aplicable a curvas mixtas.

Riemann y sus sucesores decimonónicos, que solo consideraron métricas riemannianas propiamente tales, basaron en ellas sus definiciones de *geodésica* y *curvatura*. No me referiré a esta última, pero debo explicar brevemente el concepto métrico de geodésica, que fue el que Einstein tuvo presente cuando dio con la idea matriz de su teoría de la gravedad (§ 1.4). Esta es simplemente una generalización del concepto original de una curva de longitud extrema (máxima o mínima) entre dos puntos de la superficie de la tierra. Sea γ una curva parametrizada por longitud de arco que une dos puntos $\gamma(a)$ y $\gamma(b)$ en una variedad riemanniana propiamente tal $\langle M,\mathbf{g}\rangle$. γ es una *geodésica* entre $\gamma(a)$ y $\gamma(b)$ —en el sentido métrico del término— si es ya sea la más corta, ya sea la más larga de todas las curvas que unen los puntos $\gamma(a)$ y $\gamma(b)$ por caminos ligeramente distintos del camino de γ; o, usando la terminología y el simbolismo del cálculo de variaciones, si la integral (32) es estacionaria, esto es, si

$$\delta\int_a^{b} \|\dot{\gamma}(u)\| \mathrm{d}u = \delta\int_a^{b} \mathrm{d}u = 0 \tag{34}$$

Levi-Civita (1917) mostró que para cada métrica riemanniana $\mathbf{g}$ sobre una variedad M hay un y solo un modo de definir el transporte paralelo a lo largo de las curvas de M que, en un sentido preciso, es compatible con $\mathbf{g}$. En el vocabulario de la § 7.12, este resultado se puede enunciar así: Para cada $\mathbf{g}$ sobre M hay una y solo una conexión

lineal sin torsión ∇ sobre M tal que $\nabla\mathbf{g} = 0$. Esta conexión ∇ se llama la *conexión de Levi-Civita* de **g**. La métrica **g** y la conexión ∇ corresponden al mismo tensor de curvatura sobre M. Si **g** es una métrica riemanniana propiamente tal hay perfecta coincidencia entre sus geodésicas (es decir, las curvas de longitud extrema) y las geodésicas de ∇ (es decir, las curvas cuyas tangentes son paralelas entre sí a lo largo de la curva respectiva). Si **g** no cumple la condición **M3** —esto es, si **g** es riemanniana en mi sentido del término pero no es *propiamente* riemanniana— esta coincidencia se mantiene en el caso de las geodésicas temporaloides y espacialoides, y, mediante una ingeniosa reparametrización, puede establecerse también en cierto sentido para las geodésicas nulas. No puede haber geodésicas mixtas, pues la condición $\nabla\mathbf{g} = 0$ implica que el transporte paralelo no afecte el valor de $\mathbf{g}_{\gamma(u)}(\dot{\gamma}(u),\dot{\gamma}(u))$ cuando cambia u.[25] Hasta el final de esta sección, el símbolo ∇ designa la conexión de Levi-Civita de la métrica riemanniana que estemos considerando.

Si $\langle M,\mathbf{g}\rangle$ es una variedad riemanniana y x es una carta definida sobre $U_x \subseteq M$, entonces, la métrica **g** es igual sobre U_x a una combinación lineal de los covectores $\mathrm{d}x^k$ $(1 \leq k \leq n)$:

$$\mathbf{g} = g_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j \qquad (35)$$

Los campos escalares g_{ij}: $U_x \rightarrow \mathbb{R}$ son los *componentes métricos*. Adviértase que, aunque el número de ellos es n^2, debido a la simetría de **g**, $g_{ij} = g_{ji}$, de modo que no más de $\sum_{r=1}^{n} r$ pueden ser diferentes entre sí. Los componentes g_{ij} se llaman a veces componentes *covariantes* de **g**, para contrastarlos con sus componentes *contravariantes* g^{ij}: $U_x \rightarrow \mathbb{R}$, que se definen en función de los otros por las siguientes relaciones:

$$g^{ik}g_{kj} = \delta^i_j \qquad (36)$$[26]

Si el lector ha estudiado algo de álgebra lineal recordará que toda función bilineal está asociada a una determinada forma cuadrática.[27] En nuestro caso, si llamamos $\mathbf{q}_p$ a la forma cuadrática asociada a la función bilineal $\mathbf{g}_p$ que la métrica **g** asigna al punto $p \in M$, ella está definida por

$$\mathbf{q}_p(v) = \mathbf{g}_p(v, v) \tag{37}$$

para cada vector $v \in \mathrm{T}_pM$. En vista de esto, podemos relacionar unívocamente con la métrica **g** un objeto **q**, que asigna, por un lado, a cada $p \in M$ la forma cuadrática $\mathbf{q}_p$ definida por la ecuación (37), mientras que, por otro lado, asocia a cada campo vectorial $V \in \mathcal{V}^1(M)$ el campo escalar $\mathbf{q}(V): p \mapsto \mathbf{q}_p(V_p)$. Es oportuno observar que la aplicación $\mathbf{q}: \mathcal{V}^1(M) \to \mathcal{F}(M)$ no es una aplicación lineal del módulo $\mathcal{V}^1(M)$ en su cuerpo de escalares $\mathcal{F}(M)$, pues, si $V \in \mathcal{V}^1(M)$ y $f \in \mathcal{F}(M)$, entonces $\mathbf{q}(fV) = f^2\mathbf{q}(V)$.

Este objeto **q** asociado unívocamente a una métrica riemanniana **g** es igual al cuadrado de lo que en la literatura clásica de la geometría diferencial se conoce como *elemento de línea* de la métrica (simbolizado ds). Relativamente a una carta x,

$$\mathbf{q} = ds^2 = g_{ij}dx^idx^j \tag{38}$$

La ecuación (38) se confunde fácilmente con la ecuación (35). Sin embargo, el simbolo dx^idx^j en la ecuación (38) no es una manera anticuada de escribir el producto tensorial que hoy simbolizamos así: $\mathrm{d}x^i \otimes \mathrm{d}x^j$. El objeto designado con dx^idx^j no es un campo tensorial, pues su valor en un punto $p \in M$ no es una función bilineal, sino una función cuadrática sobre T_pM, a saber, $v \mapsto \mathrm{d}x^i{}_p(v)\mathrm{d}x^j{}_p(v)$, que, en virtud de la definición (4), equivale a $v \mapsto (vx^i)(vx^j)$, para cada $v \in \mathrm{T}_pM$.[28]

Los componentes g_{ij} de la métrica riemanniana **g** relativos a una carta x y los componentes Γ^i_{jk}, relativos a la misma carta, de su conexión de Levi-Civita ∇, definidos en las ecuaciones (18), satisfacen las siguientes ecuaciones:

$$\Gamma^i_{jk} = \tfrac{1}{2}g^{ih}\left(\frac{\partial g_{hj}}{\partial x^k} + \frac{\partial g_{kh}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^h}\right) \tag{39}$$

Las ecuaciones (39) le confieren un significado métrico a los componentes $R^i{}_{jkh}$ del tensor de curvatura R presentados en las ecuaciones (25). Obsérvese que, si reemplazamos los Γ^i_{jk} en las ecuaciones (25) por las expresiones equivalentes conforme a las ecuaciones (39), en las ecuaciones resultantes entran las derivadas de segundo orden de

los componentes métricos respecto a las coordenadas. Como dije a continuación de (25), en el vocabulario de Einstein, la contracción $C^1{}_4(R)$ es el *tensor de Ricci*. En términos de los componentes métricos, sus componentes relativos a una carta x pueden expresarse como se indica a continuación:

$$\begin{aligned} R_{ij} = R^k_{ijk} &= \frac{\partial \Gamma^k_{ik}}{\partial x^j} - \frac{\partial \Gamma^k_{ij}}{\partial x^k} + \Gamma^s_{ik}\Gamma^k_{sj} - \Gamma^s_{ij}\Gamma^k_{sk} = \\ &= \frac{\partial^2}{\partial x^i \partial x^j}\left(\log\sqrt{|g|}\right) - \frac{\partial \Gamma^k_{ij}}{\partial x^k} + \Gamma^s_{ik}\Gamma^k_{sj} - \Gamma^s_{ij}\frac{\partial}{\partial x^s}\left(\log\sqrt{|g|}\right) \end{aligned} \tag{40}$$

donde g designa el determinante de la matriz (g_{ij}).[29] Comparando la primera línea de las ecuaciones (40) con la observación que hice a continuación de la (26), el lector verificará que el tensor de Ricci es un tensor simétrico si ∇ es una conexión sin torsión (como, por definición, ocurre aquí donde ∇ designa una conexión de Levi-Civita).

Me referiré ahora a las operaciones sobre tensores y campos tensoriales que se llaman "subir y bajar índices". Espero que la explicación siguiente aclare tanto su sentido y alcance como el origen de su nombre pintoresco. Consideremos una vez más la variedad riemanniana $\langle M,\mathbf{g}\rangle$. Cualquier campo tensorial T de tipo (q,r) sobre una subvariedad abierta de M puede asociarse a un campo tensorial $T^{\maltese}$ con el mismo rango y dominio que T, pero que es (i) contravariante en determinado índice en que T es covariante o (ii) covariante en un determinado índice en que T es covariante. En el caso (i) pasamos de T a $T^{\maltese}$ "subiendo el índice" en cuestión. Esto se hace formando el producto tensorial $\mathbf{g} \otimes T$ y contrayendo. Específicamente, si es en el el k-ésimo índice que T es contravariante T y $T^{\maltese}$ es covariante, $T^{\maltese} = C_2{}^{k+2}(\mathbf{g} \otimes T)$, la contracción de $\mathbf{g} \otimes T$ en el segundo índice (en que por cierto es covariante) y el $(k+2)$-ésimo (que tiene que ser contravariante, como el k-ésimo índice de T). Para ver esto más claramente supongamos que T es un campo tensorial de tipo (1,3) y que buscamos el campo tensorial de tipo (0,4) asociado a T. Los componentes T_{ijkh} de este campo tensorial relativos a una carta idónea, están dados

por las ecuaciones

$$T_{ijkh} = g_{is} T^{s}_{\ jkh} \tag{41}$$

De un modo análogo, en el caso (ii) se pasa de de T a $T^{\maltese}$ "bajando índices", para lo cual contraemos el producto tensorial **g*** ⊗ T de la métrica contravariante **g*** —definida en la nota 24— por T. Los componentes de **g*** son precisamente los g^{ij} introducidos en las ecuaciones (36). Las ecuaciones

$$T^{ij}_{\ \ kh} = g^{is} T^{i}_{\ skh} \tag{42}$$

dan entonces los componentes del campo tensorial de tipo (2,2) asociado a nuestro campo tensorial T de tipo (1,3). La asociación entre campos tensoriales del mismo rango pero diverso tipo así generada por la métrica es unívoca. Por tanto, habitualmente se considera a los diversos campos tensoriales asociados mutuamente por las operaciones de "subir y bajar índices" como *distintas formas* del *mismo* objeto. Uno se refiere entonces a los $T^{ij}_{\ \ kh}$ como componentes de la forma de tipo (2,2) de T y a los T_{ijkh} como los componentes de su forma totalmente covariante. Estos términos se aplican también a los tensores en un punto.

Para terminar, definiré el *escalar de curvatura* de la variedad riemanniana ⟨M,**g**⟩. Consideremos los componentes R_{ij} del tensor de Ricci introducidos en las ecuaciones (40). Entonces, los componentes de su forma de tipo (1,1) están dados por

$$R^{k}_{j} = g^{ki} R_{ij} \tag{43}$$

El escalar de curvatura es el campo escalar R que se obtiene contrayendo en sus dos índices el tensor de Ricci de tipo (1,1). En términos de componentes relativos a una carta,

$$R = R^{i}_{\ i} \tag{44}$$

8

Suplementos

8.1 El disco giratorio

En varios escritos posteriores al descubrimiento de las ecuaciones de campo de la relatividad general, Einstein habla de la geometría no euclídea que supuestamente se constataría mediante varas de medir rígidas en un disco que gira con velocidad angular uniforme en el espaciotiempo minkowskiano (Einstein 1916*e*, pp. 774–775; 1917*a*, pp. 53–56; 1922*c*, pp. 38–39; también ahora en CP, vol. 7, pp. 270–272, parte de un manuscrito póstumo fechado tentativamente en 1920). En el primero de los escritos citados la única conclusión que Einstein extrae de su análisis de este caso se refiere a la imposibilidad de definir coordenadas que midan objetivamente distancias y tiempos sobre el disco (1916*e*, pp. 775–776). Sin embargo, tanto en el escrito de 1922 como en el manuscrito póstumo mencionado, Einstein sugiere francamente que el disco giratorio suministra un argumento directo a favor de la curvatura del espacio. Me parece que tal sugerencia no es aceptable en el ámbito de la relatividad general, pues la curvatura constante e igual a cero *del espaciotiempo* minkowskiano no es alterada en lo más mínimo por la mera presencia de un disco que gira, cuya masa no se menciona y cabe suponer insignificante, y la curvatura *del espacio* depende de como se elija la hipersuperficie espacialoide a que esa curvatura corresponde. Pero nada impide que la reflexión sobre el disco giratorio, que figura desde 1909 en la correspondencia de Einstein[1] y asoma en un artículo suyo de 1912,[2] haya efectivamente pesado en su ánimo cuando decidió que una teoría de la gravedad tenía que basarse en una geometria espaciotemporal de curvatura variable, determinada por la distribución de la materia. En tal caso, es legítimo calificar la geometría del disco de "eslabón perdido" en la evolución de la antigua a la nueva geometría física,

como hizo John Stachel (1980). La siguiente nota, parafraseada de mi libro *Relatividad y geometría* (Torretti 1996, pp. 147–149), ayudará al lector a formarse su propia opinión sobre este asunto.

Consideremos una región del universo donde no se perciben efectos de la gravedad. Sea $\mathcal{R}_1$ un marco inercial de referencia y sea $\mathcal{R}_2$ otro marco de referencia no inercial que gira con velocidad angular constante $\boldsymbol{\omega}$ en torno a un punto fijo O de $\mathcal{R}_1$. Llamaré $\mathcal{S}_k$ al espacio relativo de $\mathcal{R}_k$, esto es, al conjunto de las posiciones espaciales que permanecen en reposo con respecto a $\mathcal{R}_k$ ($k = 1, 2$). Las relaciones métricas entre los puntos de $\mathcal{S}_k$ se establecen mediante varas de medir rígidas en reposo en $\mathcal{R}_k$. Suponemos que cada vara $\mathfrak{m}$ concuerda en cada instante con una vara estándar en reposo en el marco inercial en que $\mathfrak{m}$ reposa en ese instante. (Esto implica que las varas que reposan en $\mathcal{R}_2$ no son afectadas por la aceleración, o que los efectos de esta —por ejemplo, la tensión longitudinal en las varas que apuntan perpendicularmente hacia el eje de rotación— se corrigen antes de utilizar los resultados de las mediciones.)[3] Llamo $\mathbf{G}_k$ a la geometría determinada así mediante tales varas de medir en el espacio $\mathcal{S}_k$. Si la relatividad especial vale en la región exenta de gravedad que estamos considerando, $\mathbf{G}_1$ es la geometría euclídea. En tal caso, como ahora veremos, $\mathbf{G}_2$ es una geometría no euclídea. Sea C_1 un círculo en $\mathcal{S}_1$ con centro en O y radio perpendicular al eje de rotación de $\mathcal{R}_2$ e igual a r varas. Postulamos que $r|\omega| < c$, donde c es la velocidad de la luz en $\mathcal{R}_1$, expresada en varas por unidad de tiempo. Sin embargo, r tiene que ser lo bastante grande —o la vara estándar lo bastante pequeña— para que la circunferencia de C_1 se cubra ceñidamente con $2\pi r$ varas, ordenadas de modo que se toquen los extremos de cada par de varas consecutivas. Sea C_2 el lugar geométrico de todos los puntos de $\mathcal{S}_2$ que en un momento dado coinciden con puntos de C_1. Por simetría, C_2 tiene que ser un círculo, el mismo en todo momento (de suerte que no tenemos que ocuparnos con la definición del tiempo en $\mathcal{R}_2$). Calcularemos el cociente entre la circunferencia y el radio de C_2 en la geometría $\mathbf{G}_2$. Una vara que reposa en $\mathcal{R}_2$ a lo largo de un radio de C_2 tiene velocidad longitudinal 0 respecto a $\mathcal{R}_1$. Por lo tanto, la longitud de la vara en $\mathcal{R}_1$ —esto es, la distancia euclídea entre dos posiciones de sus extremos en $\mathcal{S}_1$, simultáneas en el tiempo de Einstein

adaptado a $\mathcal{R}_1$— es igual a la longitud de una vara en reposo en $\mathcal{R}_1$. En consecuencia, el radio de C_2 medido con varas fijas a lo largo del mismo es igual a *r*. Consideremos ahora una vara $\mathfrak{m}$ que reposa en $\mathcal{R}_2$ sobre la circunferencia de C_2. En cada momento, el marco inercial en que $\mathfrak{m}$ reposa en ese momento se mueve relativamente a $\mathcal{R}_2$ con velocidad $v = r|\omega|$, en la dirección de una tangente a C_1. Por lo tanto, la longitud de $\mathfrak{m}$ en $\mathcal{R}_1$ es igual a $\sqrt{1-v^2c^{-2}}$ veces la de una vara que repose en $\mathcal{R}_1$. ¿Cuántas varas en reposo sobre la circunferencia de C_2 y ordenadas de modo que se toquen los extremos de cada par de varas consecutivas cubren esa circunferencia ceñidamente? Tantas como puedan, ordenadas de ese modo, imprimir sus huellas sobre la circunferencia de C_1 simultáneamente, esto es, en un mismo instante del tiempo de Einstein adaptado a $\mathcal{R}_1$; a saber, $2\pi r\left(\sqrt{1-v^2c^{-2}}\right)^{-2}$. Por lo tanto, el cociente entre la circunferencia y el radio de C_2 es igual a $2\pi\left(\sqrt{1-v^2c^{-2}}\right)^{-2} > 2\pi$ y la geometría $\mathbf{G}_2$ es no euclídea. De aquí se sigue inmediatamente que el espacio $\mathcal{S}_2$, provisto de esta geometría, no puede aplicarse isométricamente sobre $\mathbb{R}^3$ ni describirse mediante coordenadas cartesianas.

El análisis anterior depende esencialmente de que sea posible *contar* las varas que reposan ceñidamente sobre C_2 e imprimen simultáneamente sus huellas en C_1. El siguiente supuesto de Einstein sin duda lo implica: “Entendemos que en un determinado momento *t* de $\mathcal{R}_1$ los extremos de todas las varas están determinados relativamente a $\mathcal{R}_1$” (1922*c*, p. 39; Einstein escribe *K* en vez de $\mathcal{R}_1$). Entonces, evidentemente, lo que se mide con ellas no es la circunferencia misma sino un polígono que se le aproxima. Pero sería ridículo esperar valores exactos en una geometría establecida con varas de medir. En todo caso, la longitud de C_2 en $\mathbf{G}_2$ excederá a $2\pi r$ en cualquier cantidad acordada, según $r|\omega|$ se vaya acercando a *c*. Algunos autores intentan darle a su tratamiento de este tema una apariencia de rigor expresando la longitud de la vara estándar como un diferencial *ds* que integran sobre C_2. Pero la precisión así obtenida es ilusoria y sus razonamientos no son ni un ápice más estrictos que el de Einstein, porque, como señalé en la nota 3, no es posible fijar con exactitud el marco inercial en que un cuerpo acelerado reposa en cierto instante, y esta noción

está envuelta en el cálculo de la contracción de las varas colocadas ceñidamente sobre C_2.[4]

8.2 Mellizos de distinta edad

Para aclarar el caso de los mellizos conviene hacer los cálculos correspondientes a un caso particular. El siguiente ejemplo figura en mi libro *La filosofía de la física* (Torretti 1999, pp. 277–280) con otros nombres pero con los mismos números. Digamos, pues, que Claudio y Claudia nacieron el mismo día de la misma madre en un astro que se mueve inercialmente. Claudio se queda en él toda su vida, pero Claudia es raptada a pocos días de nacer por piratas extraterrestres que la conducen en su nave espacial a una velocidad constante muy grande, digamos 0,9*c*, a otro astro, de donde retornan con ella inmediatamente al punto de partida a la misma velocidad. Cuando los hermanos se reúnen, Claudio acaba de cumplir ochenta años, pero —de acuerdo con la relatividad especial— Claudia aún no tiene treinta y cinco. Esta predicción contraría nuestras expectativas, pero nunca hemos visto a nadie que viaje a esa velocidad, y sería presuntuoso suponer que una chica que se mueve tan rápidamente envejecerá como cualquier otra. Las cosmolíneas de Claudia y Claudio no coinciden y el tiempo propio determinado según la ecuación (1.9) —en la p. 31— entre sus dos encuentros es menor a lo largo de aquella que a lo largo de esta. Por lo tanto, si los relojes biológicos de los mellizos marchan más o menos a unísono con los relojes atómicos, Claudio tiene que ser mayor que Claudia cuando ella regresa.

Evidentemente, esto no podría ser de este modo si los relojes biológicos registrasen el tiempo absoluto newtoniano; pero, hoy por hoy, nadie pretende que así sea. Los críticos de la relatividad especial la consideran absurda, no porque contradice las ideas de Newton, sino porque, según ellos, la teoría implica que Claudio cumplirá ochenta años antes que Claudia si las vidas de ambos son referidas al astro en movimiento inercial en que él reposa, y que esta cumplirá ochenta años antes que aquel si sus vidas son referidas a la nave en movi-

miento inercial en que reposa ella. Pues, según la relatividad especial, cualquier marco inercial puede ser considerado con el mismo derecho como sistema de referencia "en reposo". Este razonamiento no tiene en cuenta una diferencia física radical entre las vidas de los mellizos: mientras Claudio la pasa entera en un mismo sistema inercial, Claudia tiene que reposar por lo menos en dos sistemas inerciales distintos mientras se aleja de su astro natal a velocidad 0,9c y mientras se acerca a él a la misma velocidad pero en sentido contrario. Como el cambio de un sistema a otro demanda una aceleración, tanto más violenta cuanto más breve, en los primeros días de la relatividad especial muchos opinaron que el problema de los mellizos caía fuera del alcance de esta teoría, y tendría que resolverse con los recursos de la relatividad general. Esta opinión es errónea. Mediante la ecuación (1.9), la geometría de Minkowski le asigna una longitud a todas las curvas temporaloides, no solo a las geodésicas. La idea de que la relatividad espacial sería incapaz de evaluar la duración propia de un movimiento acelerado carece de todo fundamento.

De todos modos, nuestro problema matemático se facilita mucho si suponemos que Claudia viaja sucesivamente en dos naves espaciales en movimiento inercial entre las que salta sin dilación en el momento oportuno. La violencia del salto, que tendrá que ejecutar sin ningún entrenamiento previo, podría sin duda provocarle un envejecimiento prematuro; pero su edad cronológica, medida por los relojes atómicos que la acompañan, será en todo caso mucho menor que la de Claudio cuando vuelvan a verse. Sean x, y y z cartas de Lorentz (p. 51), adaptadas a los tres marcos inerciales que nos interesan: aquel en que reside Claudio y los dos en que Claudia reposa en su viaje de ida y en su viaje de vuelta, respectivamente. Para simplificar la aritmética, supondremos que las coordenadas de estas cartas expresan duraciones y distancias medidas en años y en años-luz. Entonces la velocidad de la luz c = 1 año-luz por año. Supongamos que Claudio permanece todo el tiempo en el origen espacial de la carta x, de modo que, para cada evento E en el curso de su vida, $x^1(\mathrm{E}) = x^2(\mathrm{E}) = x^3(\mathrm{E}) = 0$, y que Claudia se mueve en el plano $x^2 = x^3 = 0$ con velocidad 0.9. Sea P el evento de la separación de los mellizos, y Q el evento de su reunión. Designemos con J el evento a medio

camino entre P y Q, en que Claudia salta a la nave que se acerca a la cosmolínea de Claudio, esto es, a la geodésica PQ. Marquemos en PQ el evento G, simultáneo con J en el tiempo de Einstein adaptado a la nave espacial en que Claudia se aleja, y el evento H simultáneo con J en el tiempo de Einstein adaptado a la nave espacial en que Claudia se acerca; esto es, los eventos tales que $y^0(\mathrm{G}) = y^0(\mathrm{J})$ y $z^0(\mathrm{H}) = z^0(\mathrm{J})$ (Fig. 8).

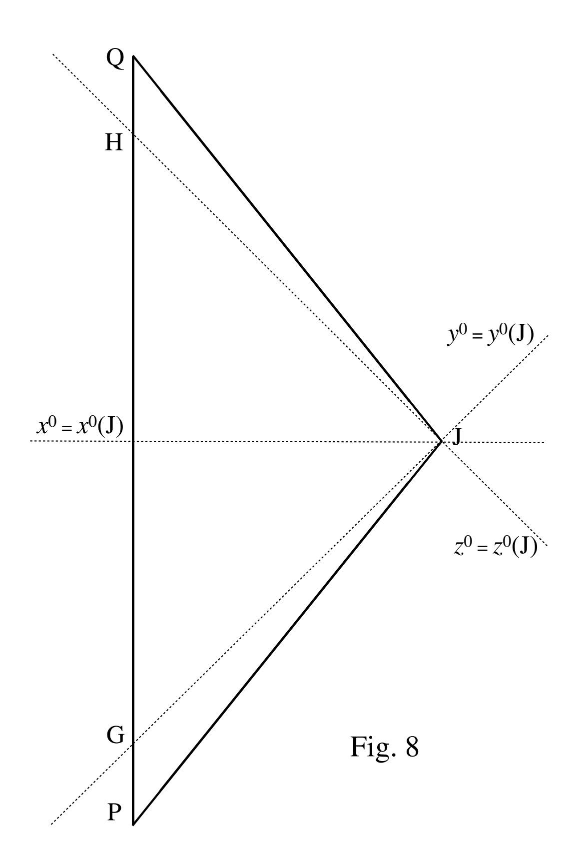


Fig. 8

Si Claudio vive ochenta años entre la partida y el regreso de Claudia, $x^0(\mathrm{Q}) - x^0(\mathrm{P}) = 80$, $x^0(\mathrm{J}) = 40$ y $x^1(\mathrm{J}) = 36$. El tiempo propio a lo largo de la cosmolínea de Claudia es igual a $(y^0(\mathrm{J}) - y^0(\mathrm{P})) + (z^0(\mathrm{Q}) - z^0(\mathrm{J}))$. Los dos sumandos se calculan fácilmente insertando los números apropiados en la primera de las ecuaciones (1.7).

$$
\begin{aligned}
y^0(\mathrm{J}) - y^0(\mathrm{P}) &= \frac{x^0(\mathrm{J}) - x^0(\mathrm{P}) - v(x^1(\mathrm{J}) - x^1(\mathrm{P}))}{\sqrt{1-(v)^2}} \\
&= \frac{40 - 0,9(36)}{\sqrt{1-(0,9)^2}} = \frac{7,6}{0,43589} = 17,4356 \\
z^0(\mathrm{Q}) - z^0(\mathrm{J}) &= \frac{x^0(\mathrm{Q}) - x^0(\mathrm{J}) - v(x^1(\mathrm{Q}) - x^1(\mathrm{J}))}{\sqrt{1-(v)^2}} \\
&= \frac{40 - 0,9(36)}{\sqrt{1-(0,9)^2}} = \frac{7,6}{0,43589} = 17,4356
\end{aligned}
\tag{1}
$$

De modo que $(y^0(\mathrm{J}) - y^0(\mathrm{P})) + (z^0(\mathrm{Q}) - z^0(\mathrm{J})) = 34,8712$, y Claudia ha vivido menos de 35 años mientras su hermano vivía 80.

Otra paradoja parece amenazarnos si admitimos que en su agenda Claudia lleve la cuenta de los cumpleaños de Claudio. Pues en el marco inercial en que Claudia reposa su hermano cumple años con menos frecuencia que ella. Así, según sus cuentas, en los 17,4356 años que ella vive entre P y J, Claudio no envejece más que 7,6 años. Y también según sus cuentas la edad de Claudio aumenta en 7,6 años entre H y Q, esto es, en los 17,4356 años que ella vive entre J y Q.[5] Por lo tanto, durante su viaje, Claudia podría celebrar a lo sumo quince cumpleaños de Claudio. Ella podría preguntarse entonces al encontrarse nuevamente con él ¿En qué etapa de mi vida completó mi hermano los otros 64,8 años que carga sobre sus hombros? Si, como es fácil suponer, Claudia es más aguda que algunos filósofos de la ciencia, seguramente se le ocurrirá pensar que, en sus circunstancias, es aconsejable usar dos cronologías, y^0 y z^0, adaptadas a los dos marcos inerciales en que sucesivamente reposa. En el tiempo y^0, los 64 cumpleaños de Claudio entre G y H están todavía por venir cuando ella

salta en J; pero en el tiempo z^0 esos 64 años ya son, en ese mismísimo instante, cosa del pasado. Claudia podría utilizar una cronología compuesta, a saber, "y^0 antes del salto y z^0 de ahí en adelante"; pero la coordenada temporal definida de este modo no es universal, ni inercial, ni se ajusta al método de Einstein, y la vida de Claudio entre G y H cae fuera de su dominio.

8.3 Fragmento sobre la idoneidad de los conceptos

Preguntábamos al final de la §3.1 ¿por qué el concepto einsteiniano de gravedad y la clasificación que impone a los fenómenos deben preferirse a las otras alternativas disponibles? ¿qué nos obliga o siquiera nos invita a pensar que este concepto se ajusta a los fenómenos que intenta captar? O dicho en otras palabras: ¿por qué pensamos que los fenómenos que la relatividad general clasifica como gravitacionales realmente van juntos y han sido bien entendidos por esa teoría? ¿hay razones que nos constriñan a pensar de este modo? Espero que el lector convendrá conmigo en que, si tales razones existen, tienen que estar afincadas en el testimonio de la experiencia.

Ahora bien, por muy optimista que uno sea con respecto al poder probatorio de la experiencia, no podrá pretender que ella nos compela a adherir perpetuamente a una teoría física particular. De hecho, la experiencia misma nos enseña todo lo contrario. Basta ver lo que ocurrió con la física desde 1788, el año en que se publicó la *Mecánica analítica* de Lagrange, hasta 1896, cuando aparecieron las *Lecciones sobre teoría de los gases* de Boltzmann; y nuevamente en el siglo trascurrido después que salió a luz el libro de Boltzmann. No obstante, el mismo ejercicio de rememoración histórica mostrará que, mientras el modo como los físicos entienden los fenómenos naturales suele cambiar drásticamente cada par de generaciones, más o menos, una vez que un grupo considerable de fenómenos de común ocurrencia ha sido reunido con éxito por una teoría, ninguna teoría posterior tratará de desagruparlos, aunque ella gire en torno a una concepción completamente distinta de esos fenómenos. Así, aunque la electro-

dinámica clásica y la cuántica conciben la radiación electromagnética de modos muy diferentes, la teoría cuántica no ha excluido de su alcance ningún fenómeno clásicamente reconocido como radiación. Esta persistencia de las clasificaciones no se debe solamente a la inercia intelectual. Volviendo a nuestro tema, me atrevería decir que, a más tardar desde el 4 de octubre de 1957, cuando el primer satélite artificial fue puesto en órbita por un misil soviético, estamos irrevocablemente comprometidos a mantener la clasificación newtoniana del movimiento planetario como caída libre. Por otra parte, no veo que nada nos constriña tan poderosa y persuasivamente a clasificar como fenómeno gravitacional el corrimiento sistemático al rojo de las líneas espectrales observado en la luz procedente de otras galaxias. Esta clasificación se adoptó porque la espectacular fuga de galaxias evidenciada a juicio de los astrónomos por el corrimiento al rojo se deducía sin más de las ecuaciones de campo de Einstein, bajo el supuesto razonable de que la materia está uniformemente distribuida a gran escala. Además, la prueba de que, según la relatividad general, un universo homogéneo se expande o se contrae, pero tiene probabilidad 0 de permanecer estático, había sido descubierta por Friedmann sin que supiera nada acerca de las mediciones del corrimiento al rojo practicadas por Hubble y sus colegas. Se disponía, pues, de una elegante explicación previa para un fenómeno que de otro modo habría sido completamente desconcertante. Pero no cabe esperar que se establezca experimentalmente una transición entre el movimiento en órbita y la expansión cósmica, por ejemplo, apilando masas a distancias apropiadamente grandes, como amañamos una transición entre la caída de los proyectiles hacia la tierra y el movimiento en órbita incrementando el empuje ascendente de un misil. Nada parece impedir, entonces, que en un remoto futuro —o en una galaxia remota— el corrimiento cosmológico de las líneas espectrales hacia el rojo reciba una explicación satisfactoria en términos no gravitacionales (e incluso no cosmológicos). Pero esta es solo una posibilidad abstracta y que no nos concierne en absoluto. La investigación actual sobre la gravedad parte del supuesto de que cualquier sustituto de la relatividad general tiene que dar cuenta de todos los fenómenos abarcados por esta teoría. Por tanto, las teorías contemporáneas concuerdan con ella

en cuanto a la extensión o denotación de 'gravedad' y disputan solo la intensión o connotación de este término.

¿Cómo llegamos a saber que la denotación acordada a un término es captada adecuadamente por su propuesta connotación? ¿cómo establecemos que una clase más o menos bien deslindada de procesos se concibe correctamente de cierta manera? Esta es quizás la pregunta más difícil de la filosofía de la ciencia y no pretendo abordarla aquí. Barrunto, sí, que a menos que nos hagamos una idea modestamente pragmática de lo que significan los adverbios 'correctamente' y 'adecuadamente', la pregunta tal vez no tenga respuesta. Por otra parte, si nos hacemos tal idea, no cabe duda de que la competencia entre la relatividad general y sus rivales contemporáneos ha sido resuelta en favor de aquella por una serie de tests experimentales. No puedo repasarlos aquí.[6] Pero querría considerar brevemente qué nos dicen sobre nuestra pregunta inicial, "¿por qué hay que concebir la gravedad como curvatura del espaciotiempo?"

Permítaseme primero insistir majaderamente en algo que es obvio. Aunque podemos entrar en campos gravitacionales de diversa intensidad más o menos del mismo modo como entramos en zonas de buen y mal tiempo, dichos campos no presentan la curvatura local del espaciotiempo a la vista, tal como el aire frío no exhibe de suyo la velocidad media de sus moléculas. Hay que *leer* un significado geocronométrico *en* los fenómenos gravitacionales para poder medir en ellos los componentes del tensor de Riemann en este o aquel punto del espaciotiempo. Así, solo tenemos testimonios indirectos de la realidad de la curvatura espaciotemporal no nula y variable, en la forma de tests experimentales que corroboran determinados supuestos y consecuencias de la relatividad general.

Los tests más antiguos conciernen al principio de equivalencia. Los más exactos certifican la igualdad de la masa inercial y la masa gravitacional con un error inferior a 1 parte en 10^{12}. Tal igualdad, sin embargo, no equivale al principio de equivalencia de Einstein sino solo a lo que se llama "el principio de equivalencia débil" (*weak equivalence principle*). En la literatura actual se distingue además entre el principio de equivalencia "muy fuerte" (*very strong*), que enuncié en la § 3.2, y el principio "semifuerte" (*medium strong*). Este último

se refiere solo a los fenómenos no gravitacionales; según él, estos se desarrollan exactamente del mismo modo en un laboratorio inercial y en uno que cae libremente en un campo gravitacional uniforme. El principio de equivalencia semifuerte está respaldado por cada experimento que confirme la relatividad especial en una astronave. El principio muy fuerte extiende la equivalencia de los marcos de referencia a los efectos puramente gravitacionales. Si falla, la porción gravitacional de la energía que mantiene atadas las partes un cuerpo —esto es, la energía necesaria para vencer la atracción puramente gravitacional entre ellas— haría contribuciones desiguales a su masa inercial y a su masa gravitacional. En tal caso, la aceleración gravitacional de la luna y de la tierra hacia el sol serían diferentes. Esto se llama el *efecto Nordvedt*. Si es nulo, cierto parámetro será igual a 0. Los experimentos sitúan el valor de este parámetro dentro de un pequeño entorno de 0. Entre 1976 y 1994, ese entorno se ha estrechado de 0 ± 0.003 a –0.0001 ± 0.0015 (Ciufolini y Wheeler 1995, p. 115).[7]

Aunque algunos autores han sugerido lo contrario, el principio de equivalencia muy fuerte es una condición necesaria mas no suficiente de la relatividad general. En efecto, la relación entre el principio y la teoría es comparable con la que existe entre las leyes de Kepler del movimiento planetario y la ley de gravedad universal de Newton. Así como, según la teoría de Newton, las leyes de Kepler son estrictamente falsas de nuestro sistema planetario y solo valen en el caso imaginario de un sistema aislado formado por un cuerpo central finito y otro cuerpo infinitesimal que interactúa gravitacionalmente con aquel, así también, según la relatividad general, el principio de equivalencia muy fuerte solo se aplica a un campo gravitacional uniforme, que no existe en ninguna parte.

…………………………………………………………………………………[8]

Notas

Capítulo 1

[1] “Omnis enim philosophiae difficultas in eo versari videtur, ut a phaenomenis motuum investigemus vires naturae, deinde ab his viribus demonstremus phaenomena reliqua” (Newton, *Principia*, prefacio a la primera edición).

[2] Newton lo enuncia así: “Todo cuerpo persevera en su estado de reposo o de movimiento uniforme en línea recta, salvo en cuanto es compelido por fuerzas impresas a cambiar su estado” (Newton 1727, p. 13).

[3] Por el Corolario V de las leyes del movimiento, o *Principio de Relatividad de Newton*: “Los movimientos recíprocos de los cuerpos incluidos en un espacio dado son idénticos, ya sea que este espacio repose, o que se mueva uniformemente en línea recta sin rotación” (Newton 1727, p. 20).

[4] En Hall y Hall 1962, p. 104; cito el texto latino en la nota 2 al capítulo 2, p. 211.

[5] James Thomson (1884); véase el pasaje citado en la p. 46.

[6] De no ser así, sería posible infringir el principio de relatividad distinguiendo el marco de referencia en que la onda luminosa, después de un tiempo de emitida, llena una esfera, de aquellos en que llena una figura menos simétrica.

[7] Hay otro inconveniente que Einstein no menciona. Si el tiempo se define de acuerdo con este criterio en todos los puntos de un marco de referencia $\mathcal{R}$, este *no es un marco inercial* en nuestro sentido de la palabra: las partículas libres sufren un cambio discontinuo de velocidad al pasar por el punto de referencia señalado. Este asunto se considera con más detenimiento en el capítulo 2, especialmente en la nota 8, p. 212.

[8] Aunque uso ‘evento’ a veces en un sentido laxo, como sinónimo de ‘suceso’ u ‘ocurrencia’, en general empleo esta palabra en una acepción más precisa, para designar los sucesos ideales, inextensos y sin duración que las teorías de la relatividad adoptan como ingredientes elementales en su descripción de los fenómenos. Siguiendo la práctica común en la literatura de la física, llamo también ‘eventos’, por metonimia, a las localizaciones posibles de tales eventos puntuales, esto es, a los puntos del espaciotiempo.

[9] En la práctica esta suposición siempre resulta ser más o menos falsa; pero, por lo mismo, los tiempos de hecho registrados por los relojes

calzan solo aproximadamente con los previstos por la fórmula

$$t'(E^*) = t(E^*)\Big/\sqrt{1-(v/c)^2}$$

10 Relativamente a la superficie de la tierra, la Plaza de Armas de Santiago de Chile está hoy en el mismo lugar que le asignó don Pedro de Valdivia el 12 de febrero de 1541. Relativamente al marco de referencia determinado por las estrellas fijas, retorna a ese lugar una vez al año. Pero relativamente a un marco comóvil con la radiación cósmica de fondo se aleja de allí a gran velocidad desde hace casi cinco siglos.

11 En las §§ 6.3 y 6.4 (pp. 145–152), me refiero a las contribuciones de Klein y Hilbert a los fundamentos de la geometría.

12 Para avanzar sin tropiezos hacia la noción de espaciotiempo, especialmente como está concebida en la relatividad general (§ 1.4), tenemos que suponer que las coordenadas de un evento son números puros (cf. la caracterización de las variedades n-dimensionales en las pp. 165–166). Por lo tanto, cuando digo que t representa el tiempo medido en segundos quiero decir simplemente que, para cualquier evento E, $t(E)$ es un número real igual a la *cantidad* de segundos trascurridos entre E y el evento O tal que $t(O) = 0$, si O precede a E, e igual a esa misma cantidad multiplicada por -1 si E precede a O. Elijo las unidades del Sistema Internacional porque son familiares para todos; pero lo que importa es que los dos sistemas comparados utilicen la misma unidad de tiempo y la misma unidad de distancia. Si las unidades utilizadas fuesen diferentes tendríamos que introducir factores de escala en las ecuaciones de transformación de coordenadas.

13 Vale decir, hay un evento O tal que $t(E) = t'(E) = x(E) = x'(E) = y(E) = y'(E) = z(E) = z'(E) = 0$.

14 Obviamente, si $\mathcal{R}'$ se moviese con velocidad constante en una dirección arbitraria la transformación cinemática dependería de tres parámetros v_x, v_y y v_z, representativos de los componentes de su velocidad en la dirección de los tres ejes del otro triedro sin primas. Pero no es necesario considerar este caso entre las transformaciones simples, pues se lo puede obtener combinando una transformación cinemática dependiente de un solo parámetro —como la representanda por las ecuaciones (5)— con dos rotaciones que lleven el eje de las x y el eje de las x' a apuntar en la dirección del movimiento.

15 Un *grupo* puede definirse como un conjunto G de objetos cualesquiera, que incluye un objeto distinguido e, y sobre el cual se ha definido una

operación $*$ que asigna a cada par de objetos $x, y \in G$ un objeto $x * y \in G$ y cumple las condiciones siguientes para cualesquiera $x, y, z \in G$:
(i) $x * (y * z) = (x * y) * z$ (la operación del grupo es *asociativa*),
(ii) $x * e = e * x = x$ (e se comporta como el cero en la adición de enteros: es el elemento *neutro* del grupo), y
(iii) existe un elemento $x' \in G$ tal que $x * x' = x' * x = e$ (cada elemento tiene un *inverso*).

En un *grupo de transformaciones*, los objetos del conjunto G son transformaciones, esto es, funciones biyectivas (uno-a-uno) que aplican un conjunto K sobre sí mismo (en nuestro caso, $K = \mathbb{R}^4$, el continuo que forman los cuádruplos de números reales $\langle t,x,y,z \rangle$); la operación del grupo es la composición de funciones; el objeto neutro e es la transformación idéntica, que asigna cada cuádruplo $\langle t,x,y,z \rangle$ a sí mismo). El grupo G_3 *generado por* dos grupos G_1 y G_2 de transformaciones de K es simplemente aquel cuyos elementos son todas las transformaciones que pertenecen a uno de los dos grupos o resultan de la composición de transformaciones pertenecientes a cualquiera de los dos. Más precisamente: $g \in G_3$ si y solo si $g \in G_1$ o $g \in G_2$ o hay objetos $h, k \in G_3$ tales que $g = h * k$.

16 Véase Einstein (1905*r*), pp. 897–902 o cualquier manual de teoría de la relatividad especial. Explico la derivación de Einstein y otras derivaciones alternativas en Torretti 1996, pp. 56–82.

17 En 1983 la Conferencia Internacional de Pesos y Medidas definió el metro como una determinada fracción de la distancia recorrida por una señal luminosa en un segundo (véase la p. 52). En virtud de esta convención, el número citado es exacto.

18 El primero en hacer presente de un modo claro e inequívoco la existencia de estos grupos fue Poincaré (1906), quien les dio el nombre de Lorentz, porque este había publicado (en 1904) la transformación (7). El adjetivo “conectado” se refiere a lo siguiente: los grupos descritos tienen la estructura de espacios topológicos conectados (cf. p. 233). Si admitiéramos la inversión del tiempo $t' = -t$, $x' = x$, $y' = y$, $z' = z$, y la transformación de paridad $t' = t$, $x' = -x$, $y' = y$, $z' = z$, obtendríamos los grupos completos de transformaciones de Lorentz (entre sistemas de coordenadas con un origen común) y de Poincaré (que incluye las traslaciones), los cuales constan de varios componentes topológicamente desconectados. Pero la inversión temporal y la paridad quedaron excluidas por nuestra convención inicial sobre la dirección en que crece cada una de las cuatro coordenadas de un sistema.

[19] En la nota 15 caractericé a $\mathbb{R}^4$ como el continuo que forman todos los cuádruplos de números reales. Normalmente se sobreentiende que hereda su topología de la recta real $\mathbb{R}$ según la regla siguiente: el producto cartesiano de cualesquiera cuatro intervalos abiertos $(a_i,b_i) \in \mathbb{R}$ $(1 \leq i \leq 4)$ es un abierto de $\mathbb{R}^4$. Pero la topología así obtenida puede también caracterizarse a partir de la geometría minkowskiana: es la topología más gruesa en que todas las transformaciones del grupo conectado de Poincaré son funciones continuas.

[20] "Um nirgends eine gähnende Leere zu lassen, wollen wir uns vorstellen, daß aller Orten und zu jeder Zeit etwas Wahrnehmbares vorhanden ist" ("Para que en ninguna parte bostece un vacío, vamos a figurarnos que en todo lugar y en todo momento está presente algo perceptible"— Minkowski 1909, p. 104).

[21] En la nueva notación, un número en posición elevada no es un exponente sino un índice de la variable a su izquierda. Para simbolizar potencias encerramos entre paréntesis la variable con su índice y escribimos el exponente a la derecha del último paréntesis. Así, $(x^1)^3 = x^1{\cdot}x^1{\cdot}x^1$; $(x^2)^2 = x^2{\cdot}x^2$.

[22] Obsérvese que (i) $x \circ y^{-1}$ es una función biyectiva que aplica a $\mathbb{R}^4$ sobre sí mismo; (ii) $y^{-1}: \mathbb{R}^4 \to \mathcal{M}$ asigna a cada cuádruplo q de números reales el evento E tal que $y(E) = q$; (iii) por lo tanto, $x \circ y^{-1}(q) = x(E)$. En otras palabras, la biyección $x \circ y^{-1}$ sustituye el juego de cuatro coordenadas asignado a cada evento E por la carta y, con el juego de coordenadas asignado a ese mismo evento E por la carta x. No cabe duda, entonces, de que $x \circ y^{-1}$ es lo que se llama una transformación de coordenadas. Que ella pertenece al grupo de Poincaré es una consecuencia de las condiciones que impusimos a las cartas x e y.

[23] Estas definiciones suponen que hemos elegido como invariante fundamental de la geometría minkowskiana la cantidad que figura a ambos lados de la ecuación (8). Si, como suele ocurrir, hubiésemos elegido esa cantidad multiplicada por –1, habría que intercambiar los signos > y < en las definiciones de *temporaloide* y *espacialoide*. Esta elección está implícita en la elección de la llamada *signatura de la métrica* de Minkowski (véase la nota 22 al capitulo 7, p. 241). Los adjetivos *temporaloide* y *espacialoide* se pueden explicar como sigue. Si la separación entre E_1 y E_2 es temporaloide, el orden en que estos eventos se suceden es el mismo en cualquier tiempo de Einstein, no importa cual sea el marco inercial a que esté adaptado; en otras palabras, el orden temporal de los eventos con separación temporaloide es un invariante de la geometría de Minkowski. En cambio, si la separación entre E_1 y E_2 es

espacialoide, su orden de sucesión no es invariante sino que depende del marco inercial de referencia a que se adapte el tiempo de Einstein utilizado para determinar ese orden; en tal caso, hay siempre un marco inercial $\mathcal{R}$ en cuyo tiempo de Einstein E_1 y E_2 son simultáneos y $-\sigma^2(E_1,E_2)$ es igual al cuadrado de la distancia euclídea entre los lugares de $\mathcal{R}$ en que ocurren E_1 y E_2. Análogamente, la separación *nula* se llama también *luminoide* (en inglés, *lightlike*), porque solo existe entre pares de eventos tales que uno de ellos podría coincidir con la emisión y el otro con la absorción de una misma señal luminosa.

24 La analogía es imperfecta, por cierto: una recta euclídea es la más *corta* de todas las líneas que unen sus extremos.

25 Traduzco así el vocablo alemán *Weltlinie* adoptado por Minkowski; en inglés se dice *worldline*; en francés, *ligne d'univers.* De un modo análogo, podemos llamar *cosmovelocidad* de una partícula en un instante dado a la tangente a su *cosmolínea* en ese instante, y *cosmoaceleración* a la derivada de la cosmovelocidad respecto al tiempo propio. Estos términos se extienden de un modo natural a los espaciotiempos curvos de la relatividad general.

26 El siguiente razonamiento ayudará a ver que esto es así. Si p se mueve libremente conforme al principio de inercia hay un marco inercial $\mathcal{R}$ en que p reposa. Sea x una carta adaptada a $\mathcal{R}$. Si C es la cosmolínea de p entre dos eventos cualesquiera E y E', $\int_C \mathrm{d}\sigma = \int_C dx^0 = x^0(E') - x^0(E)$. Sea Γ cualquier curva temporaloide vecina que va de E a E'. Entonces su longitud

$$\int_\Gamma \mathrm{d}\sigma = \int_\Gamma \sqrt{(dx^0)^2 - c^{-2}\left((dx^1)^2 + (dx^2)^2 + (dx^3)^2\right)}\,.$$

Como la expresión $(dx^1)^2 + (dx^2)^2 + (dx^3)^2$ no es nunca negativa y difiere de 0 al menos sobre una parte de Γ, es claro que $\int_C \mathrm{d}\sigma = x^0(E') - x^0(E) = \int_\Gamma dx^0 > \int_\Gamma \mathrm{d}\sigma$, de modo que C es una geodésica.

27 Einstein y Laub (1908*a*, p. 532) ni siquiera le reconocen este mérito. A propósito de la presentación de las ecuaciones fundamentales de la electrodinámica de los cuerpos en movimiento por Minkowski (1908), dicen lo siguiente: "Como este trabajo impone al lector exigencias matemáticas bastante grandes, no nos parece superfluo derivar esas ecuaciones por una vía elemental, que por lo demás concuerda esencialmente con la seguida por Minkowski". El esfuerzo invertido en ello recuerda la presentación de la mecánica celeste en los *Principia* de

Newton, sin emplear el cálculo. Cf. asimismo las penosas observaciones de Laub sobre Minkowski y uno de sus admiradores en la carta que dirigió a Einstein el 18 de mayo de 1908 (Einstein CP, vol. 5, pp. 119–120). En un manuscrito póstumo sobre la relatividad especial que empezó a escribir en 1912, Einstein volvió a tratar buena parte de lo que había elaborado con Laub, usando esta vez el formalismo de Minkowski (Einstein, CP, vol. 4, documento 1).

28 En julio de 1910, Einstein le escribe a Sommerfeld, a propósito del primero de estos trabajos: "Su nuevo artículo me causa una alegría extraordinaria. ¿Cómo puede usted pensar que yo no sabría apreciar la belleza de una investigación así? La consideración de las relaciones formales en cuatro dimensiones me parece un progreso comparable a la introducción de funciones complejas en la hidrodinámica y la electrostática de dos dimensiones. Probablemente no me expresé bien a este respecto en Salzburgo delante de usted [en una reunión celebrada en septiembre de 1909—R.T.]. Las condiciones del acontecer (ecuaciones diferenciales) son simétricas en cuatro dimensiones; este conocimiento facilita el descubrimiento de esas condiciones. El límite de la significación del enfoque cuadridimensional está, me parece, en que la cuatro dimensiones no figuran del mismo modo en las *soluciones* de esas ecuaciones que nos interesan." (Einstein, CP, vol. 5, p. 246.)

29 Cito el relato de Einstein en la p. 72. A él se refiere mi artículo "Una idea feliz", en Torretti (1994), pp. 51–65.

30 Newton (1727, p. 400) dice haber verificado experimentalmente esta relación con un error inferior a $1{:}10^3$. En el siglo XX este margen de error se fue estrechando: Eötvös, Pekar y Fekete (1922) lo redujeron a $5{:}10^9$; Roll, Krotkov y Dicke (1964) a $1{:}10^{11}$, y Braginsky y Panov (1972) a $1{:}10^{12}$. Es, pues, uno de los resultados mejor confirmados de la física.

31 En el prólogo a la traducción checa de su libro de divulgación (1917*a*). No soy capaz de leer esta traducción y el manuscrito alemán del prólogo aún no ha aparecido en los *Collected Papers*, pero he visto el pasaje pertinente citado por John Stachel (1980), p. 12.

32 "…das System der Relativitätstheorie ungemein vereinfachenden Hilfsbegriffe…" (Einstein, CP, vol. 6, p. 69.)

33 Según las explicaciones que Einstein ofrece antes sobre lo que entiende por "un sistema de referencia admisible" (*ein berechtigtes Bezugssystem*), x, y, z son coordenadas cartesianas definidas sobre un marco inercial K y $u = ict$, donde t es un tiempo de Einstein adaptado a K, c es la velocidad de la luz en el vacío e $i = \sqrt{-1}$. El uso de números imaginarios para

la coordenada temporal permite darle al invariante fundamental (8) de la geometría de Minkowski una formulación que a primera vista se distingue de la distancia euclídea (6) solo por el número de dimensiones en juego (cuatro en vez de tres).

34 El potencial gravitacional es un campo escalar Φ (tal como en la teoría newtoniana reformulada por Poisson). Entonces $c = c_0(1 + \Phi c^{-2})$, donde c_0 es una constante (Einstein 1911*h*, p. 906). Este artículo, entregado a *Annalen der Physik* el 21 de junio de 1911, apareció el 1 de septiembre; el 10 de agosto Einstein había escrito a Laub: "El tratamiento de la gravitación desde el punto de vista de la teoría de la relatividad presenta serias dificultades. Creo probable que el principio de la constancia de la velocidad de la luz en su versión habitual valga solo para espacios de potencial gravitacional constante" (CP, vol. 5, p. 309).

35 En carta a Ehrenfest del 12 de febrero Einstein dice que "la teoría de Abraham es enteramente insostenible" (CP, vol. 5, p. 408). Refiriéndose a la versión modificada (Abraham, 1912*g*), escribe el 29 de octubre a Sommerfeld: "Hasta donde veo, la nueva teoría de Abraham es lógicamente correcta, pero es solo un aborto de la perplejidad (*eine Missgeburt der Verlegenheit*)" (CP, vol. 5, p. 505). Por esta fecha, Einstein, con la ayuda de Gauss y de Riemann, ya había superado su propia perplejidad. ¿Habría calificado su teoría del campo gravitacional estático con la misma severidad?

36 "Si referimos los sistemas físicos a un sistema de coordenadas *K* en términos del cual las leyes naturales toman la forma más simple posible, hay infinitos sistemas de coordenadas referidos a los cuales esas leyes son las mismas, a saber, todos los sistemas de coordenadas en movimiento de translación uniforme respecto a *K*" (Einstein, 1912*h*, p. 1059).

37 "Pronto vi que al introducir transformaciones no lineales como lo requiere el principio de equivalencia la simple interpretación física de las coordenadas tiene que perderse; vale decir, ya no puede exigirse que las diferencias de coordenadas expresen los resultados inmediatos de mediciones con varas de medir y relojes ideales. Este conocimiento me afligía mucho, pues por largo tiempo no pude comprender qué habrían de significar entonces las coordenadas en la física." (Einstein, 1934, pp. 252–253). Einstein vuelve sobre este tema en sus "Notas autobiográficas". Ya en 1908 había comprendido las limitaciones de la teoría especial. "¿Por qué entonces se necesitaron siete años más para establecer la teoría general de la relatividad? La razón principal consiste en que uno no se libra tan fácilmente de la idea de que hay que atribuir a las coordenadas un significado métrico inmediato" (Einstein, 1949, p. 66). Entre las

reflexiones que contribuyeron a esta liberación podría contarse quizás la relativa al cociente entre la circunferencia y el radio de un disco en rotación que Einstein presenta en varios escritos y que según Stachel "brinda al parecer un 'eslabón perdido" ('*missing link*') en la cadena de razonamientos que lo condujo a la idea decisiva de que se requería una métrica no plana para un tratamiento relativista del campo gravitacional" (1980, p. 1). Me refiero a la geometría del disco giratorio en el suplemento §8.1, pp. 191–194.

38 El simbolismo oculta una diferencia más profunda: mientras que en la ecuación (13), la forma cuadrática $\sum_{i=1}^{2}\sum_{j=1}^{2} g_{ij}du^i du^j \geq 0$ siempre, la que aparece bajo el radical en la ecuación (15) podría ser negativa si la curva *C* fuese espacialoide o mixta; pero el supuesto inicial de que *C* es una cosmolínea excluye esta posibilidad en el caso considerado.

39 Conviene subrayar que esta ley geodésica de la caída libre, de enorme valor heurístico en el desarrollo de la relatividad general, no es un principio de la teoría. En efecto, las ecuaciones de campo de Einstein (1915*i*) no son lineales y es de suponer que en un sistema cerrado, no expuesto a influencias externas arbitrariamente especificables, las cosmolíneas de las partículas estarán constreñidas por su interacción gravitacional con arreglo a esas ecuaciones. Cabe preguntarse entonces si tales constreñimientos son compatibles o no con una ley que asigna una cosmolínea geodésica a cada partícula en caída libre. No se conoce una respuesta general a esta pregunta y quizás no exista, pero en diversos casos se ha establecido exactamente o mediante aproximaciones que la ley geodésica —aplicada a partículas eléctricamente neutras y sin spin— es una *consecuencia necesaria* de las ecuaciones de campo. En particular, es geodésica la cosmolínea de una partícula de prueba, esto es, de cuerpo tan pequeño que su contribución al campo gravitacional no influye significativamente sobre su propio movimiento (Geroch y Jang, 1975). También se puede probar que lo son las cosmolíneas de la materia cuando esta consiste en polvo incoherente (como en los modelos cosmológicos de Friedmann). Whittaker (1928) demostró que las señales luminosas describen geodésicas nulas.

40 Véase especialmente Norton (1985, 1984), Stachel (1989), Earman y Glymour (1978*a*,*b*). Vizgin (2002) estudia la influencia de las ideas de Mach que no menciono aquí por falta de espacio. Cf. también mi propio relato: Torretti (1996), capítulo 5.

41 El término 'objeto geométrico' fue introducido como sinónimo de 'invariante' por Schouten y van Kampen (1930) y adoptado por Veblen

y Whitehead en su enjundiosa e influyente monografía *The Foundations of Differential Geometry* (1932), donde lo explican así: "Cualquier cosa que no es alterada por transformaciones de coordenadas se llama un invariante. Así, un punto es un invariante y también lo es una curva o un sistema de curvas. También, estrictamente hablando, cualquier cosa, como una planta o un animal, que no esté relacionada con el espacio de que hablamos es un invariante. Para un invariante relacionado con el espacio […] usaremos asimismo el término *objeto geométrico*" (p. 46).

Capítulo 2

1 Hay una literatura abundantísima sobre este tema. Menciono algunos títulos en la bibliografía de Torretti (1996). Los siguientes escritos, no mencionados allí, me resultaron muy útiles cuando redacté este capítulo: Redhead (1993), Debs y Redhead (1996); Earman (1995), capítulos 6 y 7.

2 Newton, *De gravitatione et aequipondio fluidorum*, en Hall y Hall (1962), p. 104: "Unumquodque durationis momentum […] per universa spatia, suo more, sine aliquo partium ejus conceptu diffundi intelligimus."

3 Paráfrasis mía del texto de Lorentz 1895 (en Lorentz, CP, vol. 5, p. 84). Whittaker (1951/53), vol. I, pp. 406s. da una breve demostración del teorema de los estados correspondientes. En 1899, Lorentz reformula el teorema casi en los mismos términos, pero da una definición diferente de las coordenadas con acento; véase Lorentz, CP, vol. 5, pp. 149s.; cf. Torretti (1996), p. 292, n. 26.

4 ¿Leyó Einstein el artículo de Poincaré antes de escribir el suyo? No lo sé. "La mesure du temps" apareció en 1898 en una conocida revista francesa *de filosofía* y Poincaré luego lo incluyó en su segundo libro filosófico, *La valeur de la science*, publicado en 1904. Sabemos que Einstein leyó y discutió el primer libro filosófico de Poincaré, *La science et l'hypothèse*, junto con sus amigos Habicht y Besso en algún momento entre 1902 y 1905.

5 Newton enunció este principio como corolario V de sus leyes del movimiento:

> *Los movimientos recíprocos de los cuerpos incluidos en un espacio dado son idénticos, ya sea que este espacio repose, o que se mueva*

uniformemente en línea recta sin rotación.

(Newton 1726, p. 20)

Este principio de relatividad de Newton es atribuido comúnmente a Galileo, que observó en un pasaje muy citado que los fenómenos del movimiento dentro de la cabina de una nave no pueden revelarnos si la nave está en reposo o se mueve suavemente sobre un mar tranquilo (EN VII, 212-214). Sin embargo, un rasgo esencial de la segunda alternativa propuesta por Galileo —según él la entendía— era que la nave se moviera uniformemente con movimiento *circular*, sobre la superficie *curva* de la tierra.

6 Einstein lo describe como "un sistema de coordenadas en el que valen las ecuaciones mecánicas de Newton" (1905*r*, p. 892); pero a fin de cuentas resulta que una sola de las leyes newtonianas del movimiento rige inalterada en el marco considerado por Einstein, a saber, la primera, el principio de inercia: *Cada cuerpo persiste en su estado de reposo o de movimiento uniforme en línea recta, a menos que sea compelido a cambiar ese estado por fuerzas impresas sobre él.*

7 Sarkar y Stachel (1999) se toman en serio este método einsteiniano para definir la simultaneidad a distancia y lo aducen para refutar el teorema de Malament (1977) según el cual la simultaneidad definida por el método *estándar* de Einstein (explicado en la p. 51) es la *única* equivalencia entre eventos ocurridos en distintos puntos de un marco inercial $\mathcal{R}$ que (*a*) puede definirse puramente en términos de su mutua *conectibilidad causal* y (*b*) no es *universal* (vale decir, hay más de una clase de equivalencia), ni *trivial* (vale decir, hay más de un evento en cada una). Sin embargo, como muestro en la nota 8, la definición chistosa de Einstein deja sin efecto el principio de inercia, cuya validez sin embargo es indispensable para que el marco de referencia adoptado pertenezca a la clase prescrita por Einstein (véase la nota 6).

8 Para comprobarlo, imagínese una partícula libre que se mueve relativamente a $\mathcal{R}$ por un camino recto que pasa por *A*. Señálense en esta ruta una serie de puntos a cada lado de *A*, separados por intervalos de igual longitud. Supongamos que al pasar por cada punto la partícula envía señales luminosas hacia *A*. Mientras la partícula va acercándose estas señales llegan a *A* con una frecuencia mayor que cuando ya se va alejando. Si el tiempo se define por el método propuesto en broma por Einstein, el tiempo en que cada señal es emitida es exactamente el mismo en que es recibida en *A*. Por lo tanto, mientras se acerca, la partícula recorre las distancias iguales que separan a cada punto del siguiente en

menos tiempo que mientras se aleja. Al pasar por *A* la velocidad constante de la partícula disminuye bruscamente, tal como el tono percibido de una sirena policial cae de repente cuando el vehículo en que está instalada nos pasa de largo.

9 El concepto de ‘carta de Lorentz’ definido aquí es más general que el concepto de ‘carta estándar de Lorentz’ definido en la p. 30. Como se recordará, este último solo se aplica a cartas cuyas coordenadas, orientadas por la derecha, miden tiempos en segundos y distancias en metros.

10 Tal convención forzaría, por cierto, a los matemáticos a modificar su definición de espacio métrico mediante nuevos axiomas mucho más complicados que los que usan actualmente, y compelería a los guardacostas a recalibrar cada media hora los instrumentos que emplean para determinar los límites de las aguas territoriales de su país (si ε, como cabe suponer, está referido al marco inercial de las estrellas fijas). Pero el filósofo convencionalista aspira solo a que *le reconozcan* las libertades que reclama; no pretende *ejercerlas.*

11 Este requisito restringe drásticamente la pretendida flexibilidad de la definición de simultaneidad propuesta por Reichenbach. En su versión original, el factor ½ que aparece en la definición de Einstein (3) se reemplaza con un número real positivo arbitrario ε, menor que 1, “el cual, sin embargo, debe tener el mismo valor para todos los puntos P” (Reichenbach 1924, § 7, Def. 2). En esta cita, P designa el punto donde rebota la señal luminosa y por ende corresponde a mi punto *B*. Pero, si *t* es la coordenada temporal definida desde un punto *A* con arreglo a la convención de Reichenbach y la constante $\varepsilon \neq \frac{1}{2}$, la velocidad $\mathrm{d}\mathbf{r}/\mathrm{d}t$ de una partícula libre sufrirá un cambio brusco cuando ella pase por *A*. En una segunda versión de su propuesta (1928, § 26), Reichenbach eliminó sin explicaciones dicho requisito de constancia, permitiendo que ε sea una función continua de la dirección en que la señal luminosa viaja desde el punto de emisión hasta aquel donde rebota (ahora designado por Reichenbach con *B*). Esta segunda definición de Reichenbach fue elaborada luego por John Winnie (1970). Los lectores que la conocen han adivinado sin duda que me estaba burlando de ella cuando propuse una definición de distancia no simétrica (en el texto que precede a la llamada 10, en la p. 52).

12 Véanse las referencias indicadas en la nota 16 del capítulo I (p. 205).

13 Ello inspiró el siguiente sofisma deslumbrante. Supongamos que al mediodía del 18 de septiembre estoy sentado en el marco inercial $\mathcal{R}$. Sea *A* mi estado actual y sean *B* y ¬*B* dos estados míos incompatibles entre sí, uno de los cuales se cumplirá la próxima Navidad si tomo cierta

decisión el 19 de septiembre. Sea *C* un evento simultáneo con *A* en $\mathcal{R}$. Si *C* ocurre lo bastante lejos de mí, hay un marco inercial $\mathcal{R}'$ en que *C* es simultáneo con la próxima Navidad y, por ende, con *B* o $\neg B$. Como *C* está ocurriendo ahora, la decisión que tomaré o no en el día de mañana tiene que estar determinada ya. Para un contraargumento con referencias a las publicaciones originales, véase Stein (1991).

14 Esta predicción de la relatividad especial fue comprobada por Rossi y Hall (1941) observando las párticulas muy veloces y de corta duración que producen los rayos cósmicos al penetrar la atmósfera terrestre. Constataron que la vida media de estas partículas, medida con relojes en reposo en la tierra, *es mayor que lo que sería si ellas se produjeran en un laboratorio terrestre* y concuerda bien con la ecuación (5).

15 Un test de esta predicción de la relatividad especial está implícito en el experimento de Hafele y Keating (1972) con un juego de relojes atómicos que volaron alrededor del mundo en aviones comerciales. El principal propósito de este experimento era comprobar la predicción de la relatividad general sobre el comportamiento de los relojes a distinta altura (donde el potencial gravitacional es diferente).

16 Para los experimentos mecánicos, esta conclusión se deduce de la identidad entre masa inercial y gravitacional y el corolario VI a las leyes del movimiento de Newton:

> *Si cuerpos que se mueven unos respecto a otros de cualquier modo son urgidos en la dirección de líneas paralelas por fuerzas aceleradoras iguales, seguirán moviéndose unos respecto a otros del mismo modo que si no hubieran sido urgidos por esas fuerzas.*
>
> (Newton 1726, p. 21)

Einstein extiende este principio newtoniano de equivalencia a toda clase de experimentos, como hizo con el principio newtoniano de relatividad. Su audaz determinación ha sido respaldada por la experiencia hasta ahora.

17 En la acepción precisa de ‘evento’ definida en la nota 8 al capítulo 1 (p. 203).

18 Cf. Einstein, Infeld y Hoffmann (1938), p. 65; Einstein e Infeld (1949), p. 209.

19 Hawking (1969) demostró que un espaciotiempo de la relatividad general admite una partición en tajadas o lonchas tridimensionales espacialoides si y solo si no contiene curvas cerradas temporaloides o nulas y una pequeña perturbación de la métrica no puede generar ninguna curva de este tipo. Evidentemente, un tiempo universal entraña una partición como la descrita.

[20] Si nuestro planeta estuviese situado en un espaciotiempo gödeliano, todas las galaxias aparecerían en un sector del cielo y no homogéneamente distribuidas, como —a gran escala— se ven ahora.

[21] Que ningún modelo de la relatividad general puede ser una representación perfectamente adecuada de nuestro universo se deduce del hecho de que la teoría no es compatible con el valor finito de la constante de Planck h.

[22] Aprovecho la oportunidad de decir un par de cosas sobre la expresión "viajar en el tiempo", que suele utilizarse para describir lo que haría una persona cuya vida efectivamente traza una curva temporaloide cerrada. En el uso corriente, 'viajar' alude a traslados en el espacio. y la expresión 'viajar en el tiempo' es metafórica. H. G. Wells la empleó para referirse a la actividad fantástica de trasladarse a voluntad a cualquier época pasada o futura. Pero nada de eso puede ocurrir sobre una curva temporaloide cerrada, pues no hay sitio allí para una voluntad, ni un pasado, ni un futuro, ni lo que propiamente llamamos 'tiempo'. Por lo tanto, me atrevería a decir que una curva temporaloide cerrada es justamente una clase de cosmolínea a lo largo de la cual *no se puede viajar en el tiempo*. Me parece claro, en cambio, que un traslado en el tiempo está ocurriendo sin cesar justamente a lo largo de las cosmolíneas abiertas. Sin embargo, aunque seamos arrastrados —como dijo Hölderlin— "de hora en hora, como agua arrojada de barranco en barranco" incluso mientras permanecemos encerrados en una celda, no llamamos a esto 'viajar en el tiempo' porque, en contraste con los viajes regulares en el espacio, no podemos apurar el paso a nuestro gusto, ni detenernos en un instante hermoso o que nos llame la atención.

[23] Sobre las soluciones de Friedmann, véase el capítulo 5.

[24] En su recensión del libro *The Arguments of Time*, donde este capítulo se publicó por primera vez, Richard Healey (2002, p. 459) reconoce que esto resolvería lindamente la parte temporal de la primera antinomia kantiana de la razón pura, "siempre que podamos confiar en la relatividad general en las condiciones extremas del universo muy temprano". Comparto plenamente la reserva expresada por Healey; como sugiero en esta misma página (nota 21), y hago explícito en las pp. 99 y 223 (nota 24), en esas condiciones extremas la teoría no merece crédito.

Capítulo 3

[1] Mi primer bosquejo del artículo en que se basa este capítulo incluía una sección sobre la cuestión de idoneidad. El fragmento reproducido en versión castellana en la §8.3 (pp. 198 y ss.) iba a ser el comienzo de esa sección. Sin embargo, cuando vi que la cosa no se dejaba tratar en el espacio de que disponía, detuve la escritura y reemplacé ese tema con los tratados en la §3.3.

[2] Poincaré (1905) resume este artículo en cuatro páginas, presentadas a la Academia de Ciencias de París pocos días antes de que el trabajo de Einstein (1905*r*) fuese enviado a la revista *Annalen der Physik.*

[3] Einstein desechó la teoría del 11 de noviembre el día 25 del mismo mes, cuando publicó las ECE (1915*i*). Con todo, estas ecuaciones dan la misma solución de la anomalía de Mercurio, debido a que concuerdan exactamente con las ecuaciones del 11 de noviembre en el vacío interestelar en que circulan los planetas.

[4] No está demás recordar que, cuando —muchos años más tarde— alguien publicó resultados que contradecían la predicción de la relatividad general sobre el avance del perihelio de Mercurio, Einstein escribió a Born: “Si no se conociera ninguna desviación de la luz, ni avance del perihelio, ni corrimiento de las líneas espectrales, las ecuaciones de la gravitación serían, sin embargo, convincentes, pues evitan el sistema inercial (ese fantasma que actúa sobre todo, sin que las cosas actúen de vuelta sobre él). Es de veras notable que los hombres sean generalmente sordos a los argumentos más poderosos, mientras que siempre tienden a sobrestimar la precisión de las mediciones” (carta de 12.05.1952; Einstein, Born y Born 1969, p. 258).

[5] Como el astrónomo A. Hall que propuso reemplazar el factor $1/r^2$ en la ley de Newton con $1/r^{2,00000016}$ (*Astronomical Journal*, **14** (1894), 49; debo esta referencia a W. Rindler).

[6] Nótese que el término ‘carga’, que la física actual aplica por antonomasia a la eléctrica, adquirió esta acepción técnica porque la carga eléctrica desempeña en la ley de Coulomb exactamente el mismo papel que la susceptibilidad a gravitar en la ley de gravedad de Newton. En la conversación ordinaria, ‘carga’ a secas sigue siendo la que llevan los burros.

[7] Este resultado de Eötvös fue cuestionado; pero Eötvös, Pekar y Fekete (1922) lo validaron y mejoraron ligeramente; este segundo resultado fue comunicado a la Academia de Budapest en 1909 y Einstein puede haberlo conocido entonces. Posteriormente, el resultado nulo de Newton y Eötvös

fue confirmado por Roll, Krotkov y Dicke (1964) con un error de $1{:}10^{11}$ y por Braginsky and Panov (1972) con un error de $1{:}10^{12}$.

8 Al explicar su principio de equivalencia a la reunión de Viena mencionada en la p. 69, Einstein hizo este comentario: "El experimento de Eötvös juega en este respecto un papel similar al del experimento de Michelson en relación con el problema de la detectabilidad física de un movimiento *uniforme*" (1913*c*, p. 1255).

9 En la nota 37 del capítulo 1 (p. 209), cito otro texto de Einstein sobre este asunto.

10 Me refiero al método que usa señales electromagnéticas que rebotan, esto es, señales de radar. Está descrito en las pp. 19–20 y 51.

11 Podemos, por ejemplo, elegir el metro como unidad de longitud y hacer la unidad de tiempo igual a 1/299.792.458 s; o elegir el segundo como unidad de tiempo y el segundo-luz como unidad de longitud (1 segundo-luz es igual a la distancia que la luz recorre en el vacío en un segundo, esto es, 299.792.458 m, exactamente).

12 El concepto de *paralelismo a lo largo de una curva* se explica en la §7.12, pp. 176 y ss.

13 Véase en la p. 36, el pasaje que sigue a la llamada 31, así como la nota correspondiente, en la p. 208.

14 Sobre el adjetivo 'riemanniana' usado aquí para calificar la métrica del espaciotiempo —en vez de 'peudoriemanniana' o 'semiriemanniana', como se dice comúnmente— véase la nota 21 del capitulo 7 (p. 241). Sobre la signatura de una métrica riemanniana, véase la nota 22 del mismo capitulo (ibid.).

15 Véase la ecuación (7.28), en la p. 182 y las referencias indicadas en la nota 20 del capítulo 7 (p. 241).

16 "Todavía no he podido deducir de los fenómenos el fundamento (*ratio*) de estas propiedades de la gravedad, y yo no fabrico hipótesis (*hypotheses non fingo*)"—Newton 1726, p. 530.

17 Véanse las referencias citadas en la nota 40 del capitulo 1 (p. 210). El vicio de las ecuaciones de Einstein y Grossmann consiste en que no son generalmente covariantes, esto es, dejan de ser válidas cuando se las somete a ciertas transformaciones de coordenadas; una situación que parecerá francamente ridícula si se piensa que las coordenadas son aquí solo sistemas para identificar eventos mediante números y en principio carecen de todo significado intrínseco.

18 Doy gracias a la profesora D. G. Mayo por señalarme este trabajo poco conocido y enviarme una fotocopia del mismo.

[19] Un campo tensorial de rango 2 sobre una variedad cuatridimensional tiene, por cierto, 16 componentes g_{ij} $(0 \leq i.j \leq 3)$. Pero si es simétrico —esto es, si $g_{ij} = g_{ji}$ para cada par de índices— no habrá más de diez diferentes.

[20] Jeffreys conoce por cierto la relación formal entre el integrando *ds* definido en la ecuación (2) y el elemento lineal del espaciotiempo minkowskiano: "Cabe demostrar que la condición para que ds^2 en su forma más general pueda transformarse a la forma $c^2dt^2 - dx^2 - dy^2 - dz^2$ es que todos los 96 componentes de cierto vector —el tensor de Riemann-Christoffel— se anulen" (1919, p. 147). Pero su artículo no contiene ni la menor sugerencia de que "el tensor de Riemann-Christoffel" podría tener algo que ver con la noción geométrica de curvatura. Dicho sea de paso, el tensor a que se refiere Jeffreys tiene a lo sumo 20 componentes independientes; por otra parte, si distinguimos entre aquellos que, por simetría o antisimetría, tienen el mismo valor absoluto, la cuenta llega a 256; ignoro el origen del número 96 mencionado por Jeffreys.

[21] Vale la pena anotar que Feynman no escribe esta definición como hice en la ecuación (4), sino en términos de componentes:

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + 2h_{\mu\nu}(x)\sqrt{8\pi G} \tag{5}$$

Normalmente, esta diferencia de notación no tendría importancia; pero en este caso particular ocurre que en la literatura sobre la relatividad los $\eta_{\mu\nu}$ simbolizan los componentes de la métrica de Minkowski relativos a una carta de Lorentz (expresada en unidades tales que $c = 1$):

$$\eta_{\mu\nu} = \begin{cases} 1 & \text{si } \mu = \nu = 0 \\ -1 & \text{si } \mu = \nu = 1,\ 2 \text{ o } 3 \\ 0 & \text{si } \mu \neq \nu \end{cases} \tag{6}$$

Si los $\eta_{\mu\nu}$ tienen en la ecuación (5) su significado habitual, tenemos que suponer que también los $g_{\mu\nu}$ y $h_{\mu\nu}$ representan componentes tensoriales relativos a una carta de Lorentz. Quien sepa algo acerca del pensamiento de Einstein sobre estos asuntos no puede dejar de preguntarse como sería posible definir una carta de Lorentz sobre un campo gravitacional existente, puesto que, por definición, una carta de este tipo tiene que estar adaptada a un marco inercial y, por lo tanto, conforme al principio de equivalencia, solo puede definirse sobre un campo gravitacional si este

es globalmente uniforme. Esta embarazosa pregunta se evita utilizando la notación de la ecuación (4).

22 Por ejemplo, "identificando el punto (x^1, x^2, x^3, x^4) con el punto (x^1, x^2, x^3, $x^4 + k$), donde k es una constante, se cambia la estructura topológica de $\mathbb{R}^4$ en $\mathbb{R}^3 \times \mathbf{S}$, y se introducen curvas temporaloides cerradas en el espaciotiempo" (Hawking y Ellis 1973, p. 124).

23 Estos y otros modelos cosmológicos son el tema de los capítulos 4 y 5.

Capítulo 4

1 Aristóteles, *Metaph.* E.1, 1026[a]1-5. Cf. *Metaph.* K.7, 1064[a]23; *Physica* II.2, 194[a]6; *De Anima* III.7, 431[b]13.

2 Lo que nunca ha acabado de sorprenderme es que ella gozase y goce aún de tanto crédito entre los pensadores católicos. Supongo que esto se explica históricamente por la abrumadora superioridad intelectual de Aristóteles sobre todo cuanto podía oponerle la filosofía cristiana a principios del siglo XIII, cuando el *corpus aristotelicum* fue traducido al latín. Ya en ese entonces la escuela franciscana le hizo la guerra con denuedo, siendo derrotada más bien por la peste negra, que mató a Ockam y a muchos otros, que por la fuerza de los argumentos de sus adversarios.

3 Cohen (1980, p. 37) escribe: "[Newton] creó lo que entendía como contrapartidas puramente matemáticas de situaciones físicas simplificadas e idealizadas que más tarde podían ponerse en relación con las condiciones de la realidad reveladas por la experimentación y la observación". Según Cohen, esta "posibilidad de elaborar las consecuencias matemáticas de supuestos que están relacionados con posibles condiciones físicas sin tener que discutir la realidad física de estas condiciones en las primeras etapas es el carácter distintivo del estilo newtoniano" (p. 30).

4 Sagredo, que en los diálogos de Galileo representa el buen sentido de un caballero educado, es bien franco al respecto: "Questa così vana presunzione d'intendere il tutto non può aver principio da altro che dal non avere inteso mai nulla, perché, quando altri avesse esperimentato una volta sola a intender perfettamente una sola cosa ed avesse gustato veramente come è fatto il sapere, conoscerebbe come dell'infinità dell'altre conclusioni niuna ne intende" ("Esta presunción enteramente vana de comprender el todo solo puede provenir de no haber jamás entendido nada, porque quien hubiese ensayado una sola vez entender perfecta-

mente una sola cosa y hubiese saboreado de veras cómo está hecho el saber, sabría que, de la infinidad de las otras conclusiones, no entiende ninguna"—Galileo 1632; EN, VII, 127).

5 El dístico de Pope, ideado como epitafio, dice así:

Nature and Nature's Laws lay hid in Night.
God said, *Let Newton be!* and All was *Light.*

Copio de Westfall (1980, p. 863) la muy citada declaración de Newton a un amigo anónimo, poco antes de morir: "I don't know what I may seem to the world, but, as to myself, I seem to have been only like a boy playing on the sea shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, while the great ocean of truth lay all undiscovered before me".

6 $\mathbf{S}^3$ es la subvariedad tridimensional de $\mathbb{R}^4$ formada por todos los cuátruplos de números reales $\langle x,y,z,w\rangle$ tales que $x^2 + y^2 + z^2 + w^2 = 1$.

7 El último término del lado derecho está escrito conforme a la convención de Einstein (p. 107): $T_r^r = \sum_{r=0}^{3} T_r^r = T_0^0 + T_1^1 + T_2^2 + T_3^3$. El signo menos antepuesto al lado derecho no figura en los libros de texto actuales porque definen el tensor de Ricci R_{ij} de otra manera que Einstein; véase la §7.12 (p. 181) y nota 19 (p. 240).

8 Para formarse su propia opinión al respecto el lector puede reflexionar sobre lo siguiente. Supongamos que el cociente entre la masa inercial y la masa gravitacional de un cuerpo sea $1 \pm \varepsilon$; las mediciones experimentales realizadas por Newton antes de 1687, por Eötvös a fines del siglo XIX y a comienzos del XX, y por Braginsky y Panov hacia 1970 fijaron la cota superior de $|\varepsilon|$ en 10^{-3}, 10^{-8} y 10^{-12}, respectivamente. Basándose en tales resultados, tanto Newton como Einstein concluyeron que $\varepsilon = 0$. Por otra parte, a la luz de las observaciones del sistema planetario la cota superior de $|\lambda|$ es de unos 3×10^{-52} m^{-2}. Sin embargo, muchos cosmólogos se muestran hoy partidarios de atribuirle a λ un valor distinto de 0, no tanto por las mediciones —un tanto especulativas— de objetos lejanos que parecerían favorecerlo (cf. la nota 44 al capítulo 5, p. 235), sino porque ello fortalece sus ideas preferidas. Sobre este "retorno" de la constante cosmológica, véase Earman (2001).

9 El gran cuerpo está situado fuera del campo, presumiblemente en la singularidad que hay a todo lo largo del eje de simetría; el cuerpo pequeño se trata como partícula de prueba. La predicción correcta de la anomalía del perihelio de Mercurio se deduce de la solución de

Schwarzschild si el cuerpo grande se identifica con el sol, el pequeño con Mercurio y se prescinde de los demás planetas. Los astrónomos, claro está, no los desconocen; pero dejan a cargo de la vieja teoría newtoniana la predicción de sus efectos sobre la órbita de Mercurio.

10 Una *foliación* de una variedad *n*-dimensional $\mathcal{M}$ es una partición de $\mathcal{M}$ en subvariedades con el mismo número de dimensiones, mayor que 0 y menor que *n*. Por lo tanto, cada punto de $\mathcal{M}$ pertenece a una y solo una subvariedad de la foliación. Una hipersuperficie de $\mathcal{M}$ es una subvariedad (*n*–1)-dimensional.

11 Un espaciotiempo $\mathcal{U}$ se dice espacialmente *isotrópico* en cada punto si hay un ☆campo vectorial☆ V sobre $\mathcal{U}$ que determina en $\mathcal{U}$ una congruencia☆ $\mathcal{C}_V$ de curvas temporaloides y tiene además la propiedad siguiente: dado un punto $p \in \mathcal{U}$ y dos vectores espacialoides X_p y Y_p tangentes a $\mathcal{U}$ en p y ortogonales a V_p, hay una isometría $f\colon \mathcal{U} \to \mathcal{U}$ tal que $f(p) = p$, $f_{*p}(V_p) = V_p$, and $f_{*p}(X_p) = Y_p$. Aquí V_p es el valor del campo V en p, esto es, la tangente de p de la única curva de $\mathcal{C}_V$ que pasa por p y $f_{*p}\colon \mathrm{T}_p\mathcal{U} \to \mathrm{T}_p\mathcal{U}$ es la aplicación de "arrastre" inducida por f en $\mathrm{T}_p\mathcal{U}$. Un espaciotiempo $\mathcal{U}$ se dice espacialmente *homogéneo* si se lo puede partir en tajadas canónicas como las definidas bajo el número (i) y para cada par de puntos p y q en cada tajada canónica hay una isometría $f\colon \mathcal{U} \to \mathcal{U}$ tal que $f(p) = q$. Estas definiciones están parafraseadas de Wald (1984), pp. 92–93.

12 Sobre la propiedad de esta denominación, véase la nota 1 al capítulo 5, p. 225.

13 Diecisiete años más tarde, cuando el telescopio de 100 pulgadas de Mount Wilson había revelado la distribución uniforme de las galaxias dentro de un radio de 10^8 parsecs, Richard Tolman escribió: "No tenemos ninguna justificación real para suponer que el universo entero tiene las mismas propiedades que la porción que ya hemos visto. Por lo tanto, aunque haremos mucho uso de modelos homogéneos en nuestras investigaciones, hay que darse cuenta de que lo hacemos primordialmente para obtener un modelo matemático definido y relativamente simple, más bien que para asegurar una correspondencia con realidades conocidas" (1934, p. 332).

14 Las soluciones de Friedmann son el tema del capítulo 5. Me ha parecido, sin embargo, que sería útil retener aquí, como una introducción concisa a este asunto, los dos párrafos que les dediqué en la versión original de este capítulo 4.

15 Entre las soluciones que cumplen las condiciones (i), (ii) y (iii), Friedmann incluye además la publicada por Willem de Sitter (1917*a*,*b*,c), que su-

pone un espaciotiempo totalmente exento de materia y por tanto hace escarnio del llamado "principio de Mach". Pero hay varios modos de entender esta solución como modelo del universo, por lo cual prefiero no considerarla aquí. Véase la §5.5, en particular la nota 27 (en la p. 231); cf. Ellis 1990, p. 100.

16 Además de las condiciones (i) y (ii), Friedmann supuso —"para simplificar los cómputos" (1922, p. 379)— que las tajadas canónicas son ortogonales a las cosmolíneas de la materia. Pero H. P. Robertson (1935, 1936) y A. G. Walker (1935, 1937) luego demostraron independientemente que esta condición *se deduce* de la condición (ii), si cada tajada canónica es esféricamente simétrica en torno al punto en que corta a cualquier cosmolínea de la materia (véanse las pp. 121 y ss.).

17 Los diversos tipos posibles de universo de Friedmann se describen en las pp. 133–134.

18 Los ☆componentes contravariantes☆ del tensor de energía de un fluido perfecto son

$$T^{ik} = (\rho + p)V^iV^k + pg^{ik}$$

donde las V^i son los componentes de la cosmovelocidad del fluido y las g^{ik} son los componentes contravariantes de la métrica. Las constantes p y ρ representan la densidad y la presión del fluido, respectivamente. Poniendo $p = 0$ obtenemos el tensor de energía del polvo incoherente, presupuesto por Friedmann.

19 Einstein (1922). Me cuesta imaginar que alguien haya creído que la seca retractación de Einstein (1923), publicada en la misma revista, bastaba para deshacer el daño causado por su primer comentario.

20 La etapa final del debate científico sobre este asunto ha sido descrito admirablemente por R. Smith (1982). Smith señala que "aunque se puede sostener correctamente que a principios de los años treinta [la teoría de los universos islas] era aceptada por casi todos los astrónomos, el contenido y la interpretación de la teoría había cambiado sustancialmente desde fines de los años diez y en el curso de los años veinte" (pp. 160–161). En efecto, el tamaño y la distancia de las nebulosas extragalácticas era sistemáticamente subestimado debido a una evaluación simplista de la relación entre período y luminosidad de las variables cefeidas. Walter Baade (1952) mostró que hay dos clases de cefeidas —asociadas con dos tipos de poblaciones estelares— en que la relación entre período y luminosidad es muy diferente. Antes del hallazgo de Baade, nuestra propia

galaxia parecía ser mucho mayor que las otras, un genuino continente entre los universos islas, en violación del principio copernicano.

21 En griego la Vía Láctea se llama ὁ γαλαξίας (*scil.* κύκλος), esto es, 'el círculo lechoso' (< γάλα = 'leche'), o quizás 'el círculo de mazamorra de trigo en leche' (ἡ γαλαξία = πόλτος κρίθινος ἐν γάλακτι—Hesych.). La voz 'galaxia', que el Diccionario de la Real Academia Española registra desde 1734 en la acepción de 'Vía Láctea' y desde 1832 como sinónimo del mineral 'galactite' (con la grafía 'galaxía', ajustada a la pronunciación griega, hasta la edición de 1927 inclusive), solo desde 1956 designa, en tercera acepción, a "cualquier otra formación estelar semejante" a la Vía Láctea. Por su parte, el *Oxford English Dictionary* documenta 'galaxy' en esta acepción con un texto de 1848 que puede traducirse así: "Soberbios grupos o galaxias separados entre sí por abismos tan tremendos que sobrepasan las distancias que dividen una estrella de otra… En medio de este sistema de cúmulos flota la galaxia cuyas glorias nos rodean más de cerca" (J. P. Nichol, *Stellar Universe*, iii. 59). En 1930, R. H. Baker escribía en inglés sin titubeos: "La existencia de galaxias (*galaxies*) más allá de la Vía Láctea (*the Milky Way*) fue finalmente demostrada por Hubble en 1925".

22 Algunas reflexiones melancólicas sobre el reconocimiento tardío de sus méritos figuran en Alpher y Herman 1990, pp. 146–151.

23 R. H. Dicke, P. J. E. Peebles, P. G. Roll y D. T. Wilkinson (1965).

24 Se acepta comúnmente que en las circunstancias actuales, las ecuaciones de campo de Einstein no pueden valer a distancias menores que la llamada longitud de Planck $l_P = \sqrt{\kappa\hbar c} \approx 10^{-35}$ m (donde κ es la constante gravitacional de Einstein que figura en las ecuaciones (1) y (2); $\hbar$ es la constante de Planck dividida por 2π, y c es la velocidad de la luz). Wald (1984, p. 379) infiere de ello que el modelo estándar tiene que fallar dondequiera que el escalar de curvatura R sea mayor que l_P^{-2}. Sin embargo, esto sería solo una cota superior para R, mas no necesariamente la mínima. La verdad es que no sabemos si la relatividad general y aun las teorías cuánticas no resultarían inaplicables a curvaturas (y densidades) mucho menores pero imposibles de reproducir en nuestros laboratorios. En ausencia de controles experimentales, cualquier conjetura sobre este asunto es pura especulación.

25 El formalismo parametrizado post-newtoniano o *formalismo PPN* permite construir modelos de los planetas y satélites en órbita en torno a cuerpos centrales, que son compatibles con distintas teorías de la gravedad. Los modelos contienen varios parámetros ajustables a la luz de las observa-

ciones. Dependiendo del valor asignado a esos parámetros el modelo es una realización de la relatividad general o de otra de esas teorías. El formalismo PNN demuestra así que, contra lo que afirman los seguidores de Thomas Kuhn, no es imposible contrastar varias teorías físicas con un mismo modelo de los fenómenos.

26 Véase también la sección titulada “Schwarzschild Solution in a Cosmological Model” en Peebles (1993, pp. 296–298). Peebles observa que, en virtud del teorema de Birkhoff, “en un modelo cosmológico homogéneo y sin presión podemos evacuar una región esférica, reemplazando el material con una masa compacta M situada en su centro sin afectar el espaciotiempo fuera de la región” (pp. 296s.). Sin embargo, no me está claro que la geometría global no resulte afectada si este procedimiento se aplica a una colección de regiones esféricas que cubre una tajada canónica del modelo, en forma de producir una distribución de masas separadas por tales regiones vacías, que coincida con la distribución actual de las estrellas; ni siquiera tengo claro que el procedimiento se deje aplicar de este modo.

27 Debo esta referencia a Misner, Thorne y Wheeler (1973); véanse las pp. 739–40 y la fig. 27.3.

28 Para las teorías de gran unificación (GUTs) la fuerza fuerte que mantiene la cohesión del núcleo atómico, la fuerza débil patente en la desintegración beta y la fuerza electromagnética familiar a todos son manifestaciones de una fuerza única, que en las circunstancias actuales del universo toma estas tres formas a primera vista irreductibles. Las GUTs tropiezan con dos dificultades no superadas aún: asignan al protón una vida media finita, larguísima pero no tanto como para que los experimentos en curso desde hace veinte años no debiesen haber registrado ya la desintegración de algunos, y predicen la existencia de monopolos magnéticos. La inflación de Guth explica por qué estos últimos no se han observado nunca.

Capítulo 5

1 Este principio no fue patrocinado por Copérnico, quien se dio por contento con trasladar el centro del universo de la tierra al sol vecino. En justicia, habría que llamarlo *el principio de Bruno*, por el fraile soñador que la iglesia romana quemó vivo el 17 de febrero de 1600. O quizás el *principio cusano* por el príncipe de la misma iglesia que hacia 1450 proclamó que el universo no tiene un centro físico ni dentro ni fuera de la tierra.

2 Friedrich Schur (1886) demostró que una ☆variedad riemanniana propiamente tal☆ que sea isotrópica en torno a cada punto también es homogénea. Para aplicar el teorema de Schur al universo sobre la base de la distribución isotrópica de la energía hay obviamente que presuponer que el universo tiene una estructura riemanniana dependiente de la distribución de la energía. Esta es una de las hipótesis fundamentales de la relatividad general.

3 Más radical aún es la postura metodológica de los partidarios de la hoy desacreditada cosmología del estado estacionario o SSC (*steady state cosmology*). Según ellos, para hacer cosmología hay que poder extraer conclusiones válidas para todo el espaciotiempo de los datos que tenemos, recogidos aquí y ahora. Por lo tanto la posición que ocupamos debe ser genérica, no privilegiada, no solo en el espacio sino también en el tiempo. Tal es el contenido del *principio cosmológico perfecto* postulado por la SSC. La expansión del universo —elemento **[2]** del modelo estándar— contradice a primera vista este principio y, para resolver la contradicción, la SSC postula la creación continua de materia, merced a la cual la densidad media del universo y por tanto su apariencia local permanecen constantes a pesar de la expansión. Más grave e irreconciliable es el conflicto entre la SSC y la radiación cósmica de fondo, cuya existencia sería incomprensible si hubiera tenido siempre la misma baja temperatura actual,

4 Sobre la voz ‘galaxia’ véase la nota 21 al capítulo 4 (p. 223).

5 Einstein mismo usó letras griegas para los índices que pueden tomar cuatro valores diferentes y latinas para los que solo toman tres, y muchísimos autores lo han imitado en esto. Siempre me ha parecido, sin embargo, que debemos utilizar el alfabeto más familiar para los índices que se refieren a las cuatro dimensiones del espaciotiempo y el más exótico para los que remiten solamente a las tres del espacio solo. Se subraya así la artificialidad de la convención que separa estas últimas.

6 En su seminario en la Universidad de Göttingen y en una conferencia sobre "El principio de relatividad" leída a la Sociedad Matemática de Göttingen el 5 de noviembre de 1907 y publicada póstumamente (Minkowski 1915). Pero Einstein probablemente se enteró de las ideas de Minkowski a través de la famosa conferencia "Espacio y tiempo" ofrecida por Minkowski a la Asamblea de Naturalistas y Médicos Alemanes el 21 de septiembre de 1908 (Minkowski 1909).

7 En 1907 Minkowski tenía solo 43 años, pero Einstein tenía 15 años menos y lo veía probablemente como una figura del pasado. A la actitud cambiante de Einstein hacia la obra de Minkowski se refieren las notas 27 y 28 del capítulo 1 (pp. 207–208).

8 Sobre el adjetivo 'riemanniana' con que califico la métrica del espaciotiempo —en vez de 'peudo-riemanniana' o 'semi-riemanniana', como se dice comúnmente— véase la nota 21 al capítulo 7 (p. 241).

9 Mi definición algebraica de vectores temporaloides y espacialoides se ajusta al uso de Einstein y de Friedmann. Hoy normalmente estos términos se definen a la inversa, de modo que $\boldsymbol{\eta}_P(v_P,v_P)$ es negativo si es temporaloide y positivo si es espacialoide. Esta diferencia depende de las diversas convenciones de signatura a que se refiere la nota 22 al capitulo 7 (p. 241). Es claro que entre esas convenciones hay que elegir una, pero no importa cuál. Lo que cuenta es que los términos 'temporaloide' y 'espacialoide' retengan el significado geométrico expresado por la condición de ortogonalidad así como las connotaciones físicas explicadas en el texto.

10 Como expliqué en la nota 8 al capítulo 1 (p. 203), llamo 'evento', en sentido propio, a los sucesos ideales, inextensos y sin duración, que la relatividad especial y la relatividad general asumen como ingredientes elementales en su descripción de los fenómenos; pero también llamo 'eventos', por metonimia, a las localizaciones posibles de tales eventos puntuales, esto es, a los puntos del espaciotiempo. A menudo, como aquí, la palabra presta buenos servicios usada en ambas acepciones a la vez. Este uso doble y a veces ambiguo es habitual en la literatura física.

11 Las partículas más rápidas que la luz o *taquiones* tendrían, por cierto, cosmolíneas espacialoides. Pero, aunque se ha escrito mucho sobre los taquiones, no hay ni el más mínimo indicio de que existan y quienes escriben sobre relatividad general normalmente suponen que no los hay.

12 Tomo la frase entre comillas de una carta de Einstein a Max Born del 12 de mayo de 1952 (Einstein, Born and Born 1969, p. 258).

13 Como dije en la p. 71, en las publicaciones en que enuncia este principio, Einstein se refiere expresamente a la equivalencia entre un labora-

torio cerrado que reposa en un campo gravitacional uniforme y otro acelerado uniformemente respecto a los marcos inerciales. Tal enunciado por cierto implica y es implicado por el que arriba doy yo.

14 En otras palabras, si $P \in \mathcal{U}$ y v y w son vectores pertenecientes a $T_P\mathcal{U}$, tales que $\mathbf{g}_P(v,v) > 0$ y $w \neq \mathbf{0}$, entonces $\mathbf{g}_P(v,w) = 0$ solo si $\mathbf{g}_P(w,w) < 0$. Una métrica riemanniana que cumpla esta condición suele llamarse una *métrica de Lorentz* (aunque en los escritos de Lorentz no se hace nunca una alusión a ella).

15 Expresada en términos de la constante gravitacional de Newton G = 6,673(10) × 10^{-11} $\mathrm{m^3 kg^{-1} s^{-2}}$ y la velocidad de la luz c = 299.792.458 m/s, $\kappa = 8\pi G/c^4$.

16 El DRAE incluyó por primera vez en la edición 2001 la voz **cizalladura**, definida así: "Deformación producida en un sólido por la acción de dos fuerzas opuestas, iguales y paralelas". En Chile los ingenieros usan desde hace tiempo la voz ‘cizalle’ para designar la acción deformante de tales fuerzas. En francés se dice ‘cizaille’, en inglés, ‘shear’.

17 Esta es la versión original de las ECE, publicada el 25 de noviembre de 1915 (Einstein 1915*i*):

$$R_{ij} = -\kappa(T_{ij} - \tfrac{1}{2}g_{ij}T) \qquad (2)$$

donde $T = T^k_{\ k}$ es el escalar formado por contracción☆ del tensor de energía respecto a sus dos índices. Einstein (1917*b*) formuló la versión corregida así:

$$R_{ij} - \lambda g_{ij} = -\kappa(T_{ij} - \tfrac{1}{2}g_{ij}T) \qquad (3)$$

Probar la equivalencia de la formulación (1) en la p. 110 y esta formulación (3) es un ejercicio simple de álgebra tensorial. La formulación (1) se ve más natural, por cuanto el tensor de la energía, que debe estar dado (o postularse), figura mondo y lirondo en el lado derecho y los componentes g_{ij} —que son las incógnitas del sistema de ecuaciones— aparecen confinados al lado izquierdo. Sin embargo, en la generalidad de los casos los g_{ij} estarán implícitos también en el tensor de energía (sin ir más lejos, ya en el caso sencillísimo de un fluido perfecto en las soluciones cosmológicas de Lemaître; véase la nota 18 al capítulo 4, p. 222). Además, con $\lambda = 0$, la versión (2) es preferible en las aplicaciones a la mecánica celeste: en el vacío interestelar en que orbitan los planetas, el tensor de energía es igual a 0 y las ecuaciones (2) se reducen simplemente a $R_{ij} = 0$. En todas mis formulaciones de las ECE el lado derecho va precedido de un signo menos. Esto se debe a que —en aras del interés histórico de este trabajo— me atengo a la definición del tensor de

Ricci adoptada por Einstein y respetada por Friedmann. La diferencia entre esta definición y la habitual ahora se explica en la §7.12 (p. 181 y nota 19, p. 241), pero la repito aquí: hoy formamos el tensor de Ricci por contracción del tensor de Riemann en los índices primero y tercero; pero Einstein aplicaba la contracción a los índices primero y cuarto.

18 Schwarzschild era director del observatorio astrofísico de Postdam. A la sazón, sin embargo, servía a su país en el frente oriental, donde pocos meses más tarde contrajo la enfermedad infecciosa que le causó la muerte. Schwarzschild presentó su solución de las ECE como un perfeccionamiento de la solución aproximada publicada por Einstein el 18 de noviembre 1915 (1915*h*). En vista del debate filosófico sobre la infradeterminación de la teoría por la experiencia, vale la pena señalar que este trabajo de Einstein no pretendía ser una solución de las ecuaciones (2), que publicaría una semana más tarde, sino del siguiente sistema de ecuaciones publicado la semana anterior (Einstein 1915*g*).

$$R_{ij} = -\kappa T_{ij} \tag{4}$$

Las ecuaciones (4), propias de una teoría de la gravedad que Einstein pronto desechó por considerarla físicamente inverosímil, concuerdan con las ecuaciones (2) dondequiera que el tensor de energía $T_{ij} = 0$. De este modo, las confirmaciones empíricas de la solución de Schwarzschild (1916*a*) dan el mismo respaldo a dos teorías conceptualmente muy distintas (Einstein 1915*g* y Einstein 1915*i*). La solución de Schwarzschild fue descubierta independientemente por Droste (1916), quien introdujo las coordenadas polares que ahora se utilizan habitualmente para presentarla.

19 Digo que las coordenadas (x^0, x^1, x^2, x^3) que usa Schwarzschild son supuestamente globales porque (*a*) se extienden hasta el infinito; (*b*) Schwarzschild dice que las coordenadas espaciales x^1, x^2, x^3 son "ortogonales" (*rechtwinklig*), lo que normalmente significa en alemán que son cartesianas, y (*c*) él no nos advierte expresamente que hay puntos del espaciotiempo que sus coordenadas no cubren. Hoy sabemos que cualquier modelo de la solución de Schwarzschild para el espacio vacío (1916*a*) incluye una variedad tridimensional de puntos en esta situación. Cada modelo se caracteriza por una constante de integración, igual a dos veces la masa *m* atribuible al cuerpo central. El factor de uno de los términos del elemento de línea de Schwarzschild es el valor recíproco de la diferencia entre 2*m* y la suma de los cuadrados de las coordenadas espaciales. Si esta suma es igual a 2*m*, ese término conlleva una inadmisible división por 0. Por lo tanto, las coordenadas x^1, x^2 y x^3 no pue-

den estar definidas en ningún punto p del campo de Schwarszchild donde, si lo estuvieran, $(x^1(p))^2 + (x^1(p))^2 + (x^1(p))^2$ sería igual a $2m$.

20 En verdad, en un espaciotiempo $\langle \mathcal{U},\mathbf{g}\rangle$ que satisfaga las condiciones de la solución de Schwarzschild para el espacio vacío, la métrica **g** no está definida en el origen de las coordenadas espaciales utilizadas por Schwarzschild (ni en las usadas por Droste; cf. nota 18). No hay nada en $\mathcal{U}$ que corresponda a la recta en $\mathbb{R}^4$ formada por los cúadruplos $\{\langle t,0,0,0\rangle: t \in \mathbb{R}\}$, y esto no puede remediarse mediante una transformación de coordenadas. Para ver qué significa esto, consideremos una secuencia $p_1, p_2,\dots$ de puntos de $\mathcal{U}$, cuyas coordenadas espaciales convergen al triple $\langle 0,0,0\rangle$. Los físicos dicen comúnmente que el campo métrico **g** tiene una singularidad en el punto al cual converge esta secuencia; pero este modo de expresarse, tomado de teorías cuyos campos de fuerzas están definidos sobre una variedad diferenciable cuya métrica no depende de ellos, puede causar confusión. La secuencia $p_1, p_2,\dots$ no converge a un punto del dominio del sistema de coordenadas x ni converge a un punto del espaciotiempo perteneciente al dominio de otro sistema de coordenadas. Por lo tanto, en el espaciotiempo $\langle \mathcal{U},\mathbf{g}\rangle$ *no hay ningún punto* al que la secuencia $p_1, p_2,\dots$ pueda converger. Por lo tanto, el campo gravitacional axisimétrico de la solución de Schwarzschild para el espacio vacío *no tiene sitio* para el cuerpo central masivo que supuestamente le da sustento. Schwarzschild (1916*b*) publicó sin embargo otra solución, llamada *interior*, que también es estática y axisimétrica como la primera, pero contempla la presencia de un fluido perfecto en una región finita que rodea —e incluye— el eje de simetría. Esta región es el lugar ocupado por el cuerpo central. La solución interior de Schwarzschild supone que el fluido en cuestión tiene densidad uniforme y por lo tanto es incompresible. Aunque esta idealización es común en la hidrodinámica clásica, es inadmisible en la teoría de Einstein pues implica que el sonido se propaga instantáneamente a través del fluido. Sin embargo, los físicos han sabido eludir este obstáculo (véase, por ejemplo, Misner, Thorne y Wheeler 1973, pp. 609s.) y han utilizado con provecho la solución interior de Schwarzschild en el estudio del colapso gravitacional de una estrella.

21 Cf. el pasaje de una carta de Einstein a de Sitter citado en la nota 25.

22 El concepto de *foliación* de una variedad n-dimensional fue definido en la nota 10 al capítulo 4 (p. 221). Aquí hablamos de una partición de tal variedad en hipersuperficies, esto es, en subvariedades $(n-1)$-dimensionales (tridimensionales si la variedad en cuestión es un espaciotiempo).

Cada punto de la variedad pertenece a una y solo una hipersuperficie de la foliación.

23 Estos conceptos se definieron en la nota 11 al capítulo 4 (p. 221). La condición E2 significa que, para cada $p \in \mathcal{U}$, el grupo de las rotaciones SO(3) actúa transitivamente como grupo de isotropía de p sobre la hipersuperficie de la foliación en que p está situado.

24 $\mathbf{S}^4$ es la subvariedad cuatridimensional de $\mathbb{R}^5$ formada por todos los quíntuplos de números reales $\langle x,y,z,w,u \rangle$ tales que $x^2 + y^2 + z^2 + w^2 + u^2 = 1$; es, pues, el análogo en cinco dimensiones de una superficie esférica incrustada en el espacio tridimensional ordinario. En cambio, el universo de Einstein (1917*b*) es homeomorfo a $\mathbb{R} \times \mathbf{S}^3$, esto es, a una hipersuperficie *cilíndrica* en $\mathbb{R}^5$. (La superficie de un cilindro ordinario en el espacio euclidiano de tres dimensiones es homeomorfa a $\mathbb{R} \times \mathbf{S} \subset \mathbb{R}^3$.)

25 Einstein CP, vol. 8, p. 422; énfasis de Einstein. De acuerdo con la convención adoptada en la p. 107, reemplazo los índices griegos de Einstein por letras de nuestro alfabeto. W. de Sitter (1917*a*, p. 1225) cita el texto alemán de este pasaje. Einstein prosigue: "Mientras no se cumplió ese postulado, para mí no se había logrado el propósito de la teoría general de la relatividad. Ello solo vino a obtenerse gracias al término en λ." La observación final resulta bastante curiosa si, como dice Einstein (1917*b*, p. 152, citado arriba en la p. 116), dicho término se introdujo exclusivamente para asegurar "una distribución cuasi-estática de la materia" y no se requería para eliminar las condiciones de frontera que él reputaba incompatibles con el principio de Mach.

26 Las dos últimas citas provienen de un *postscriptum* añadido en octubre de 1917 a de Sitter 1917*c*. En un trabajo comunicado a la academia de Amsterdam el 30 de junio de 1917, de Sitter ya había señalado que a un observador que viva en su universo esférico "las líneas en los espectros de los objetos muy distantes tienen que aparecerle desplazadas hacia el rojo" (1917*b*, p. 235). Cita allí las grandes velocidades radiales positivas atribuidas a NGC 4594 y NGC 1068 por Pease, Slipher and Moore, así como la velocidad radial menor, pero así y todo considerable, que exhibe la galaxia de Andrómeda según Wright, Pease y Slipher (p. 236 n.1). Años más tarde, Tolman menciona otra razón para preferir la solución de W. de Sitter a la de Einstein: "La solución exterior de Schwarzschild [...] para el campo estático en el espacio vacío que rodea una masa esférica de materia converge al elemento de línea de de Sitter para un universo completamente vacío si permitimos que la masa m de la esfera de materia en el origen tienda a cero" (1934, p. 349).

27 De hecho, el universo de de W. Sitter solo exhibe cumplimiento de la condición E3 si se lo representa en lo que Schrödinger (1956, p. 14) llama "el marco estático". Este conlleva una coordenada temporal cuyas curvas paramétricas no son geodésicas y no concuerdan con el tiempo propio a lo largo de la cosmolínea de una partícula de prueba en caída libre. En una representación más natural, el universo de W. de Sitter resulta ser el caso límite de un universo de Friedmann en expansión cuando la densidad media tiende a cero en todas partes. Robertson y Noonan lo describen como "único modelo estacionario no-estático" (1968, p. 365), donde el epíteto "estacionario" se refiere a un modelo cosmológico que "presenta a los observadores el mismo aspecto en *distintos momentos*" (p. 346).

28 Como explico en las pp. 188–189, un campo tensorial dado de rango r sobre una variedad riemanniana puede ser visto como un campo tensorial covariante de tipo $(0,r)$, como un campo tensorial contravariante de tipo $(r,0)$, o como un campo tensorial mixto de tipo (p,q), para cualquier par de enteros $\langle p,q \rangle$ tales que $p+q = r$. Estas formas diversas pueden reducirse unas a otras formando sus productos tensoriales con formas apropiadas de la métrica y operando contracciones. En aras de la brevedad, llamo componentes *covariantes* y *contravariantes* de un campo tensorial a los componentes de la forma covariante y de la forma contravariante de ese campo tensorial, respectivamente.

29 Según los índices i y j recorren el conjunto $\{0,1,2,3\}$ se obtienen 16 pares de índices; pero, debido a que todos los tensores en juego son simétricos, hay seis pares redundantes.

30 Si las cosmolíneas de la materia forman una congruencia $\mathcal{C}$ en el espaciotiempo $\mathcal{U}$ hay un campo vectorial V: $\mathcal{U} \rightarrow \mathrm{T}\mathcal{U}$ que asigna a cada punto $p \in \mathcal{U}$ la tangente en p a la única curva de $\mathcal{C}$ que pasa por p. Los componentes covariantes y contravariantes de V satisfacen la ecuación $V_i V^i = 1$. Como el vector V_p es la cosmovelocidad en p de la partícula material cuya cosmolínea es esa curva, es natural llamar *cosmovelocidad de la materia* al campo vectorial V.

31 En vez de ello, Lemaître postuló un fluido perfecto con presión isotrópica p, dependiente del tiempo pero no del espacio. En tal caso las siguientes ecuaciones reemplazan a las ecuaciones (9):

$$T_{ij} = \rho V_i V_j + p(V_i V_j + g_{ij}) \qquad (9^*)$$

Este toque realista es significativo solo si una parte no desdeñable de la energía del universo existe en la forma de radiación. Según la física actual, ello no es así ahora, pero tuvo que ocurrir en las etapas iniciales

de nuestro universo si este se parece a un modelo cosmológico de Friedmann-Lemaître en expansión. Como mi propósito aquí es solo explicar el razonamiento matemático de Friedmann, prescindo de la presión p y me atengo a las ecuaciones (9).

32 Las condiciones G1 y G2 parafrasean el primero de los dos "supuestos de la segunda clase" referentes al “carácter geométrico del mundo”. La condición G3 es una paráfrasis del segundo. A continuación traduzco el texto de Friedmann:

> “Los supuestos de la segunda clase son los siguientes:
>
> 1. Después de asignar las coordenadas espaciales x_1, x_2, x_3, tenemos un espacio de curvatura constante, la cual, sin embargo, puede depender de la coordenada temporal x_4. Mediante la introducción de coordenadas espaciales apropiadas, el intervalo ds, determinado por $ds^2 = g_{ik}x_ix_k$, puede reducirse a la forma siguiente:
>
> $$ds^2 = R^2((dx_1)^2 + \sin^2 x_1(dx_2)^2 + \sin^2 x_1 \sin^2 x_2(dx_3)^2\ ) + 2g_{14}dx_1dx_4 + 2g_{24}dx_2dx_4 + 2g_{34}dx_3dx_4 + g_{44}dx_4{}^2.$$
>
> Aquí R depende solo de x_4. R es proporcional al radio de curvatura del espacio, el cual por tanto puede variar con el tiempo.
>
> 2. En la expresión para ds^2, g_{14}, g_{24} y g_{34} pueden hacerse iguales a 0 mediante una elección juiciosa de la coordenada temporal; dicho brevemente, el tiempo es ortogonal al espacio.”
>
> (Friedmann 1922, pp. 378–379)

33 Por lo tanto, el principio copernicano no me faculta para inferir de la isotropía observada en la distribución de las fuentes de radiación sobre mi cono de luz pretérito que *este cono* es una hipersuperficie isotrópica por doquier y por lo tanto homogénea (en virtud del teorema de Schur). Pero, basándome en el carácter genérico y no privilegiado de mi posición, puedo aseverar que cualquier otra observación real o posible ocurrirá en el vértice de un cono de luz pretérito similar, y por ello también isotrópico. Ahora bien, a menos que implícita o explícitamente someta esta aseveración a múltiples *caveats* y restricciones, no es fácil distinguirla del principio cosmológico perfecto de la SSC aludido en la nota 3 (p. 225).

34 Ulteriormente, Robertson (1935) dedujo la condición G1 de las premisas siguientes, que clarifican o quizás suplementan su postulado referente a la equivalencia de todas las perspectivas y puntos de vista: (i) los "observadores" equivalentes describen una congruencia de cosmolíneas; (ii)

dos puntos del espaciotiempo pueden unirse a lo sumo por un solo "camino de luz" que sea la trayectoria espaciotemporal de una señal luminosa directa; (iii) el camino de toda señal luminosa que un "observador" *A* pueda recibir directamente de un "observador" *B* yace entero sobre una variedad diferenciable bidimensional (*AB*); (iv) (*AB*) = (*BA*). Robertson demuestra que si se cumplen estas cuatro condiciones todos los "observadores" pueden sincronizar sus relojes y definir así una coordenada temporal global.

35 Sobre la acción de un grupo, véase la nota 12 al capítulo 7 (p. 240). Un *grupo de movimientos* es un grupo de Lie conectado de isometrías (vale decir, de transformaciones de una variedad riemanniana propiamente tal que no afectan a la métrica, sino que la dejan invariante). Un grupo *G* es un *grupo de Lie* si tiene la estructura de una variedad diferenciable y la aplicación $G \times G \to G$, $\langle x,y \rangle \mapsto xy^{-1}$ es lisa. Si el grupo de Lie *G* es una variedad *n*-dimensional, cada elemento de *G* puede identificarse mediante *n* números reales; por eso, se dice comúnmente que *G* es un grupo de *n* parámetros (o *n*-paramétrico). Como espacio topológico, *G* se dice *conectado* si no es la unión de dos abiertos no vacíos que se excluyen mutuamente. De otro modo *G* es *desconectado* y consta de dos o más subespacios conectados llamados *componentes*. El elemento neutro de *G* pertenece por cierto a uno de esos componentes y solo a uno, que constituye un subgrupo conectado de *G*. Si *G* es un grupo de isometrías de una variedad riemanniana, dicho subgrupo actúa sobre la variedad como grupo de movimientos.

36 Sean $\langle \mathfrak{H},\mathbf{g} \rangle$ y $\langle \mathfrak{H}',\mathbf{g}' \rangle$ dos variedades riemannianas propiamente tales. Una aplicación $f\colon \mathfrak{H} \to \mathfrak{H}'$ se dice *conforme* si preserva los ángulos, esto es, si, para cualquier punto $p \in \mathfrak{H}$ y cualquier par de vectores v, $w \in T_p\mathfrak{H}$, $\measuredangle(v,w) = \measuredangle(f_{*p}v, f_{*p}w)$. En vista de las ecuaciones (7.9) y (7.30) (pp. 173 y 184), es claro que esto ocurre si y solo si hay un campo escalar, positivo por doquier, $\omega\colon \mathfrak{H} \to \mathbb{R}$ tal que, para cada $p \in \mathfrak{H}$, $f^*_p\mathbf{g}'_{f(p)} = \omega(p)\mathbf{g}_p$. Llamamos a ω el *factor escalar* de la aplicación conforme *f*.

37 Robertson critica a Friedmann por no haberse dado cuenta de esto: "Aunque en el sistema de coordenadas que empleamos aquí, las cosmolíneas x^μ = const. son en efecto geodésicas, de modo que el reposo es un estado posible para el contenido material del universo, no ocurre así […] con las coordenadas de Friedmann. La condición para que esto sea admisible con coordenadas tales que las $g_{0\mu}$ sean iguales a cero es que g_{00} dependa a lo sumo de *t*, y la discusión del universo de de Sitter [por Friedmann] no cumple este requisito" (1929, p. 828, n. 9; conforme a la convención estipulada en la p. 107, he reemplazado dos veces con

la letra griega μ el índice *i* empleado por Robertson). Cf. la nota 27 (p. 231).

38 Lo adapto, sí, a la convención de signatura de Einstein y Friedmann que sigo en este libro (cf. nota 22 al capítulo 7, p. 241). De adoptarse la convención de signatura preferida hoy, habría que multiplicar por –1 los lados derechos de las ecuaciones (13), (14) y (15).

39 Conforme al uso estándar entre los científicos, digo que un universo de Friedmann *se expande* cuando las cosmolíneas geodésicas de la materia divergen y que *se contrae* cuando convergen. El lector debe tener presente, con todo, que las hipersuperficies espacialoides indexadas con t no son compactas si su curvatura constante es negativa o nula; en tal caso no se les puede asignar un volumen finito y capaz de crecer o decrecer. En el sentido ordinario de la palabra, un universo de Friedmann solo puede expandirse, esto es, aumentar de volumen, si $k = 1$. A modo de advertencia, en el texto escribí "expandido" entre comillas.

40 La ecuación (31) puede obtenerse también directamente de las ecuaciones (28) y (29). Reordenando los términos en la ecuación (28) obtenemos

$$\rho = \frac{3}{\kappa}\left(\frac{(S')^2}{S^2} + \frac{k}{S^2} - \frac{\lambda}{3}\right) \tag{28'}$$

Multiplicando ambos lados de (28′) por S^3 y diferenciando con respecto a t obtenemos

$$\begin{aligned}\frac{d}{dt}\left(\rho S^3\right) &= \frac{3}{\kappa}\frac{d}{dt}\left(S(S')^2 + kS - \tfrac{1}{3}\lambda S^3\right)\\ &= \frac{3}{\kappa}\left(2SS'S'' + (S')^3 + kS' - \lambda S'S^2\right)\\ &= \frac{3S'S^2}{\kappa}\left(\frac{2SS'' + (S')^2 + k}{S^2} - \lambda\right)\end{aligned}$$

como se comprueba fácilmente al comparar la última línea con la ecuación (29).

41 Si ρ es mucho mayor que su valor actual, no puede desdeñarse la presión de la radiación. Entonces hay que reemplazar el tensor de energía de las ecuaciones (9) por el de las ecuaciones (9*) en la nota 31 (p. 231), que representa un fluido perfecto con presión p isotrópica por doquier. Esta práctica ahora común fue iniciada por Lemaître.

42 Las figuras 4 y 5 se basan en Robertson and Noonan 1968, p. 375, figs. 17.2 y 17.1, respectivamente.

43 Si, a la luz de la ecuación (31), reemplazamos ρ con $\rho_0(S_0/S)^3$ en la ecuación (30), obtenemos

$$S'' = S^{-2}\left(\frac{\lambda S^3}{3} - \frac{1}{6}\kappa\rho_0 S_0^{\ 3}\right) = S^{-2}(aS^3 - b)$$

donde he escrito *a* y *b* en lugar de las dos constantes λ/3 y $\kappa\rho_0 S_0^{\ 3}/6$, respectivamente. Por lo tanto

$$S''' = -2S^{-3}S'(aS^3 - b) + 3aS' = S'(a + bS^{-3})$$

que es obviamente igual a 0 si $S' = 0$. La cuarta derivada de *S* respecto al tiempo es, por tanto, igual a $S''(a + 2bS^{-3}) + 6bS^{-4}(S')^2$, que también es igual a 0 si $S' = S'' = 0$. Una rápida reflexión sobre estos resultados nos persuadirá de que, para todo entero $n \geq 3$, la enésima derivada de *S* respecto al tiempo es igual a un polinomio cada uno de cuyos términos incluye como factor por lo menos una derivada de *S* respecto al tiempo de orden inferior a *n*. Por inducción se sigue entonces que $S' = S'' = 0$ implica que $d^nS/dt^n = 0$ para todo *n*.

44 Hace pocos años, tras un laborioso y refinado análisis de datos obtenidos de la observación de supernovas muy antiguas, dos grupos de astrónomos concluyeron que $S'' > 0$ (Perlmutter *et al.* 1998, Garnavich *et al.* 1998). Entonces, por la ecuación (30), $\lambda > \frac{1}{2}\kappa\rho$ y por ende, λ tiene que ser mayor que 0, a menos que $\kappa = -1$ (la menos popular de las tres alternativas).

45 Basta dar los siguientes pasos: (i) elegir una hipersuperficie $\mathfrak{H}$ de $\mathcal{U}$ en que *t* sea constante; (ii) reunir todas las cartas del atlas máximo de $\mathcal{U}$ cuyo dominio incluya un subconjunto abierto de $\mathfrak{H}$; (iii) para cada carta $x = (x^0, x^1, x^2, x^3)$ de esta colección y cada punto $p \in \mathfrak{H} \cap \mathrm{dom}\, x$, asignar las coordenadas $x^1(p)$, $x^2(p)$, $x^3(p)$, a la curva de $\mathcal{C}$ que corta $\mathfrak{H}$ en *p*. Una vez hecho esto, $\mathcal{C}$ está provista de un atlas diferenciable formado por cartas con recorrido en $\mathbb{R}^3$ y por lo tanto con la estructura de una variedad diferenciable tridimensional (cf. p. 166).

Capítulo 6

1 Cf. Locke (1690), III.xi.16, IV.iii.18, IV.xii.8.
2 Saccheri 1733, Corolario al Lema V.
3 Basándose en los datos astronómicos disponibles, "Lobachevsky halló que el defecto del triángulo formado por Sirio, Rigel y la estrella Nº 29 de Eridano era igual a $3{,}727 \times 10^{-6}$ segundos de arco, una magnitud demasiado pequeña para que fuese significativa, dado el margen de error de las observaciones" (Torretti 1978, p. 64).
4 Gauss a Bessel, 27 de enero de 1829, en Gauss WW, vol. 8, p. 200.
5 Decimos que dos rectas *a* y *b* son *coplanares* si yacen sobre un mismo plano.
6 Un conjunto *C* de números reales se dice *acotado por abajo* si hay un número real *x* que es menor que cualquier elemento de *C* excepto a lo sumo el propio *x* (si pertenece a *C*). En tal caso, se dice que *x* es una *cota inferior* de *C*. Todo conjunto de números reales acotado por abajo tiene una cota inferior máxima o *ínfimo*, esto es, una cota inferior mayor que todas las otras.
7 La carta completa está reproducida en Frege (KS), pp. 410–413; la cita procede de la p. 412. Esta discusión entre Frege y Hilbert ya había sido anticipada por Lewis Carroll (pseudónimo del matemático Charles Dodgson) en el siguiente diálogo entre su heroína Alicia y Humpty Dumpty, el protagonista de una clásica *nursery rhyme*:

> "No sé qué entiendes por 'gloria'," dijo Alicia.
>
> Humpty Dumpty sonrió despectivamente. "Claro que no lo sabes— hasta que yo te lo diga. Quise decir 'Ahí tienes un lindo argumento apabullante!'"
>
> "Pero 'gloria' no significa 'un lindo argumento apabullante'," objetó Alicia.
>
> "Cuando *yo* uso una palabra," dijo Humpty Dumpty, en un tono un tanto desdeñoso, "significa exactamente lo que yo elijo que signifique— ni más ni menos."
>
> "La cuestión," dijo Alicia, "es si *puedes* hacer que las palabras signifiquen tantas cosas diferentes."
>
> "La cuestión," dijo Humpty Dumpty, "es quién manda—eso es todo."
>
> (*Through the Looking-Glass*, en Carroll 1960, pp. 268s.)

8 Véase Torretti 1978, pp. 68–82, y arriba, § 1.4.
9 La versión original inglesa de este capítulo, escrita para la *Stanford*

Encyclopedia of Philosophy, termina con una lista de lecturas sugeridas que he incorporado a la bibliografía. Aparte de las citadas arriba o en otros capítulos, ella incluye obras de Blumenthal (1961), Boi (1995), Bonola, (1945), Freudenthal (1957, 1960), Heath (1956), Nagel (1939), Rosenfeld (1988) y Winnie (1986).

Capítulo 7

1 Como es habitual, n designa aquí —y en todo el resto de este apéndice— un entero positivo cualquiera, que se supone fijado por el contexto. Ocasionalmente se usa m para designar otro entero positivo que puede o no ser igual a n.

2 Además del concepto de espacio vectorial menciono ocasionalmente el concepto menos difundido de *módulo*. Se trata de una sencilla generalización del primero. Los escalares de un espacio vectorial pertenecen a un cuerpo (como $\mathbb{R}$ o $\mathbb{C}$); los escalares de un módulo pertenecen a un anillo. Así como hablamos de funciones lineales y multilineales que aplican espacios vectoriales en el respectivo cuerpo de escalares, hablamos de funciones lineales y multilineales sobre módulos y con valores en el respectivo anillo de escalares. Como todo cuerpo es un anillo, todo espacio vectorial es un módulo.

3 Esta definición de *aplicación lisa* se ajusta al propósito de estas notas terminológicas y, en general, al nivel de este libro. Una exposición matemática más refinada tendría que distinguir grados de lisura. En tal caso diríamos que una aplicación diferenciable f cuyo dominio es un abierto U de $\mathbb{R}^n$ y cuyo recorrido es un abierto de $\mathbb{R}^m$ (para cualesquiera enteros positivos n y m) es *una aplicación de clase* $\mathscr{C}^q$ *en un punto* $p \in U$ si f está definida en p y, para cada entero positivo $r \leq q$ y cada índice k $(1 \leq k \leq n)$, la derivada parcial de r-ésimo orden $\partial^r f(x)/\partial x_k{}^r$ existe y es continua en p. f es *una aplicación de clase* $\mathscr{C}^q$ si y solo si es de clase $\mathscr{C}^q$ en cada punto de su dominio. Obviamente, si h es un entero positivo menor que q toda aplicación de clase C^q es una aplicación de clase C^h. Cuando se usa esta terminología es común llamar lisas a las aplicaciones de clase $\mathscr{C}^1$ y *aplicaciones de clase* $\mathscr{C}^\infty$ a las que aquí yo llamo lisas. Las aplicaciones continuas suelen llamarse *aplicaciones de clase* C^0.

4 Sonaría menos inglés si dijese ‘carta n-ple’, ‘atlas n-ple’, pero estas expresiones tampoco son muy castizas. En todo caso, normalmente se habla

de *cartas* y *atlas*, sin calificativos, pues el número de dimensiones de las variedades consideradas se desprende fácilmente del contexto.

5 Sean g y h cartas $\mathcal{A}$-compatibles de S. Considérese un punto $\langle x_1,\ldots, x_n\rangle \in \mathbb{R}^n$ donde $h \circ g^{-1}$ está definida. Sea $g^{-1}(\langle x_1,\ldots, x_n\rangle) = p$. Entonces hay una carta $f \in \mathcal{A}$ definida en p. Por lo tanto, $f \circ g^{-1}$ está definida y es lisa en $\langle x_1,\ldots, x_n\rangle$ y $h \circ f^{-1}$ está definida y es lisa en $f(p)$. De modo que $h \circ g^{-1} = h \circ f^{-1} \circ f \circ g^{-1}$ es lisa en $\langle x_1,\ldots, x_n\rangle$.

6 Para demostrarlo hay que probar que cada carta en $\mathcal{A}_{\text{max}}$ es un homeomorfismo. Sea f una de esas cartas. Por definición, f es biyectiva; la topología estipulada garantiza que f es continua. Debemos probar que la aplicación inversa f^{-1} también es continua. Ahora bien, si f^{-1} no es continua hay un $p \in \text{dom } f$ que tiene un entorno abierto $U_p \subseteq \text{dom } f$ tal que $f[U_p] \subseteq \mathbb{R}^n$ no es un abierto. U_p tiene que ser el dominio de una carta g de $\mathcal{A}_{\text{max}}$, pues la topología de S es la más gruesa que garantiza la continuidad de todas las cartas de $\mathcal{A}_{\text{max}}$. Entonces $g[U_p]$ es un abierto de $\mathbb{R}^n$ y como $g \circ f^{-1}$ es lisa —y por ende continua—, $(g \circ f^{-1})^{-1}[g[U_p]]$ también es un abierto. Pero $(g \circ f^{-1})^{-1}[g[U_p]] = f \circ g^{-1}[g[U_p]] = f[U_p]$. Por consiguiente, $f[U_p]$ es un abierto y f^{-1} es continua.

7 Obsérvese que $\mathcal{F}(p)$ también es un *anillo*, con la suma definida como en la p. 167 y la multiplicación $\mathcal{F}(p) \times \mathcal{F}(p) \to \mathcal{F}(p)$, $\langle g,h\rangle \mapsto gh$ definida por $(gh)(q) = g(q)h(q)$ para cada $q \in \text{dom } g \cap \text{dom } h$. Con estas operaciones, la colección $\mathcal{F}(M)$ de todos los campos escalares definidos sobre M también es un anillo y un espacio vectorial.

8 El término *índice* refleja el simbolismo utilizado habitualmente para designar sus componentes —y también los tensores mismos, en virtud de una metonimia cuya legitimidad fue demostrada por Penrose (1968)—, esto es, una letra seguida de q índices elevados y r subíndices en determinadas posiciones, como en la fórmula (6) en la p. 172. Los términos *covariante* y *contravariante* son vestigios de una época en que los tensores eran vistos como sistemas de números que se “transforman” según ciertas reglas en virtud de una transformación de coordenadas. Einstein previno a un amigo contra este enfoque ya en 1916 (CP, vol. 8, p. 348), pero la inercia docente lo mantuvo vivo por muchos años. ‘Covariante’ significa ‘que varía junto (esto es, con las coordenadas)’, mientras que ‘contravariante’ significa ‘que varía en el sentido opuesto’. El siguiente recurso mnemotécnico puede ser útil: Un *co*vector es *co*variante en su índice único, pero solo puede figurar como argumento en las posiciones correspondientes a los índices *contra*variantes de otros tensores.

9 En esta nota explicaré cómo caracterizar con mayor propiedad la operación descrita antes de la llamada (p. 170). Definimos dos módulos sobre

el anillo $\mathcal{F}(M)$ de los campos escalares sobre M (véanse las notas 2 y 7). El módulo $\mathcal{V}^1(M)$ comprende todos los campos vectoriales sobre M y el módulo $\mathcal{V}_1(M)$ comprende todos los campos de covectores. La suma modular y la multiplicación por escalares en $\mathcal{V}^1(M)$ se definen así: Para cualesquiera campos vectoriales V, $V' \in \mathcal{V}^1(M)$, campo de covectores $\omega \in \mathcal{V}_1(M)$, campo escalar $f \in \mathcal{F}(M)$ y punto $p \in M$,

$$(V + V')_p(\omega_p) = V_p(\omega_p) + V'_p(\omega_p,)$$

$$(fV)_p(\omega_p) = f(p)V_p(\omega_p)$$

Las operaciones características del módulo $\mathcal{V}_1(M)$ se definen intercambiando 'campo vectorial' y 'campo de covectores' en el enunciado anterior. A la luz de estas definiciones es claro que el campo tensorial de tipo (q,r) considerado en el penúltimo párrafo de la p. 170 se comporta como una aplicación $(q+r)$-lineal $W: (\mathcal{V}_1(M))^q \times (\mathcal{V}^1(M))^r \to \mathcal{F}(M)$. El campo escalar $p \mapsto W_p(\omega^1{}_p,\dots,\omega^q{}_p,V^1{}_p,\dots,V^r{}_p)$ no es otra cosa que el valor de esta aplicación W en $\langle\omega^1,\dots,\omega^q,V^1,\dots,V^r\rangle \in (\mathcal{V}_1(M))^q \times (\mathcal{V}_1(M))^r$. Este resultado puede utilizarse sin más para conferirle una estructura de módulo sobre el anillo $\mathcal{F}(M)$ a la colección de todos los campos tensoriales de tipo (q,r) sobre M, que designaré con $\mathcal{V}^q{}_r(M)$. La suma modular y la multiplicación por escalares en $\mathcal{V}^q{}_r(M)$ se definen así: Para cualesquiera campos tensoriales W, $W' \in \mathcal{V}^q{}_r(M)$, campos de covectores $\omega_1,\dots,\omega_q \in \mathcal{V}_1(M)$, campos vectoriales $V_1,\dots,V_r \in \mathcal{V}^1(M)$ y campo escalar $f \in \mathcal{F}(M)$,

$$(W + W')(\omega_1,\dots,\omega_q,V_1,\dots,V_r) =$$
$$W(\omega_1,\dots,\omega_q,V_1,\dots,V_r) + W'(\omega_1,\dots,\omega_q,V_1,\dots,V_s)$$
$$(fW)(\omega_1,\dots,\omega_q,V_1,\dots,V_r) = f\cdot W(\omega_1,\dots,\omega_q,V_1,\dots,V_r)$$

donde los símbolos + y · en los lados derechos representan respectivamente la suma y la multiplicación en el anillo $\mathcal{F}(M)$.

10 Según Spivak (1979, vol. II, p. 155), "una notación estrictamente apropiada" exigiría que esta aplicación se designe con $(f_{*p})^*$. Observa que no sería posible reunir todas las $f^*{}_p$ en una sola aplicación de fibrados $f^*\colon \mathrm{T}N^* \to \mathrm{T}M^*$, pues f puede aplicar dos puntos p, $q \in M$ en un solo punto de N y $f^*{}_p$ no tiene que ser igual a $f^*{}_q$. "Por otra parte," —dice Spivak— "podemos hacer algo con el fibrado cotangente que no podríamos hacer con el fibrado tangente. Sea ω una sección de $\mathrm{T}N^*$. Entonces podemos definir así una sección η de $\mathrm{T}M^*$: $\eta(p) = \omega(f(p)) \circ f_{*p}$, es decir, $\eta(p)(X_p) = \omega(f(p))(f_{*p}X_p)$," para cada $p \in M$ y cada campo vectorial X sobre M (ibid.). "Esta sección η se designa, como es natural, con $f^*\omega$. No hay un modo correspondiente de transferir un campo vectorial X sobre M

a un campo vectorial sobre N. No obstante estas diferencias podemos decir, *grosso modo*, que una aplicación $f: M \rightarrow N$ produce una aplicación f_* en la misma dirección entre los fibrados tangentes y una aplicación f^* en la dirección opuesta entre los fibrados cotangentes" (p. 156).

11 La palabra 'geodésica' viene del griego γῆ, 'tierra' y δαίειν, 'dividir'. El prototipo de nuestras geodésicas fueron ciertas curvas estudiadas en la geodesia, rama de las matemáticas aplicadas que determina la forma de la tierra así como las figuras y áreas de grandes porciones de su superficie.

12 Para una definición de *grupo*, véase la nota 15 al capítulo 1 (pp. 204–205). Consideremos un grupo G con operación de grupo $\diamond$ y elemento neutro e. Se dice que *G actúa* sobre una variedad diferenciable M (por la izquierda) si para cada $g \in G$ hay un difeomorfismo $\rho_g: M \rightarrow M$ tal que, para todo g, $h \in G$ y todo $a \in M$, $\rho_{g \diamond h}(a) = \rho_g \circ \rho_h(a)$. Para mayor claridad, escribo ga en vez de $\rho_g(a)$. Se dice que G actúa *transitivamente* sobre M si, para cada $a,b \in S$ hay un $g \in G$ tal que $a = gb$. Se dice que G actúa *efectiva* o *fielmente* sobre M si $ga = ha$ para todo $a \in M$ solo si $g = h$. En tal caso, $ga = a$ para todo $a \in M$ solo si $g = e$. Con Göckeler y Schücker (1987, p. 119), diré que G actúa *libremente* sobre M si en el conjunto de difeomorfismos $\{\rho_g: g \in G\}$ el único que tiene un punto fijo es ρ_e, esto es, el correspondiente al elemento neutro e. En otras palabras, G actúa *libremente* sobre M si la condición $ga = a$ para un $a \in M$ implica que $g = e$. En tal caso, obviamente, la acción de G sobre M es también efectiva. El *grupo lineal general* $GL(n,\mathbb{R})$ es el grupo formado por todas las matrices reales no singulares, $n \times n$ (esto es, con n filas y n columnas), con la multiplicación de matrices como operación del grupo y la matriz diag(1,…,1) como elemento neutro.

13 Obsérvese que el dominio U_x de cualquier carta x de una variedad n-dimensional M es siempre una variedad n-dimensional paralelizable. Pues U_x es una subvariedad abierta de M y la aplicación $p \mapsto \left\langle \left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p \right\rangle$ es obviamente una sección de FU_xG.

14 Doy un bosquejo superficial en Torretti 1996, pp. 263–278. Se hallarán exposiciones más completas y mejor motivadas en Spivak 1979, vol. II, capítulo 8, y en Kobayashi y Nomizu 1963/69, vol. I, pp. 63–153.

15 En particular, la derivada covariante $\nabla_k f$ del campo escalar f en la dirección del campo vectorial $\partial/\partial x^k$ es igual a la derivada parcial $\partial f/\partial x^k$.

16 A pesar del significado infinitamente ambiguo que doy a $\dot{\gamma}$, mis definiciones de *geodésica* y *paralelismo* son inequívocas, porque dependen so-

lamente de los valores de $\dot{\gamma}$ sobre el camino de γ, que son, por cierto, los mismos como quiera que se elija el campo vectorial $\dot{\gamma}$ entre los innumerables candidatos posibles. Con todo, para quienes detestan la ambigüedad en matemáticas aunque sea inocua, Spivak explica un modo de evitarla en este caso (1979, vol. II, pp. 247–251).

17 Los lectores que piensen que la claridad requiere precisión agradecerán la elucidación siguiente. Sea dom $X \cap$ dom $Y = U \subseteq M$. Entonces $\mathcal{R}(X,Y)$ es una aplicación lineal de $\mathcal{V}^1(U)$ sobre sí mismo. Si dom $Z \neq U$, la ecuación (23) se aplica a la restricción de $\mathcal{R}(X,Y)$ a $U \cap$ dom Z.

18 Algunos autores prefieren reservar el nombre de 'tensor de Riemann' para la curvatura de una variedad riemanniana, en el sentido de la §7.13, esto es, para la curvatura generada por la conexión de Levi-Civita de la variedad. En cualquier caso, el llamado tensor de Riemann es un campo tensorial.

19 Véase Einstein, CP, vol. 6, p. 316. En la literatura actual se llama tensor de Ricci a la contracción $C^1{}_3(R)$.

20 Sobre la ecuación (28) de desviación geodésica puede consultarse el libro de Ciufolini y Wheeler (1995), pp. 31–36; hay presentaciones más concisas en las obras de Wald (1984), pp. 46–47, y Schutz, (1980), pp. 213–214. Véase asimismo Hartle (2003), pp. 450–454.

21 Casi todos los autores llaman 'semi-riemanniana' or 'pseudo-riemanniana' a una métrica que reune las propiedades **M1** y **M2** y reservan el adjetivo 'riemanniana' para aquellas métricas que poseen además la propiedad **M3**. Por difundida que esté, me resisto a imitar esta práctica, porque (i) no es habitual llamar 'semi-*X*' o 'pseudo-*X*' a un género del cual *X* es una especie, y (ii) los prefijos 'semi-' y 'pseudo-' no hacen justicia a la contribución de Riemann a la geometría física de Einstein (cuya métrica cumple las condiciones **M1** y **M2**, pero no la condición **M3**). Es verdad que Riemann mismo no contempló métricas sin la propiedad **M3**; pero también es verdad que Copérnico nunca pensó que los planetas tuvieran órbitas elípticas, no obstante lo cual nadie diría que nuestra concepción del sistema planetario es pseudocopernicana o semicopernicana.

22 Aprovecho la ocasión para explicar el concepto de *signatura* de una métrica riemanniana a que se alude en los capítulos 1, 3 y 5. Sea $\mathbf{g}$ una métrica riemanniana sobre la variedad *n*-dimensional M. Sea T_pM el espacio tangente a M en un punto p. Una base $\langle \mathbf{e}_1,\ldots,\mathbf{e}_n \rangle$ de T_pM se dice ortonormal si, para cualesquiera índices j y k, $\mathbf{g}_p(\mathbf{e}_j,\mathbf{e}_j) = \pm 1$ y $\mathbf{g}_p(\mathbf{e}_j,\mathbf{e}_k) = 0$ si $j \neq k$ ($1 \leq j,k \leq n$). Sea q el número de índices para los que $\mathbf{g}_p(\mathbf{e}_j,\mathbf{e}_j) = 1$ y r el número de índices para los que $\mathbf{g}_p(\mathbf{e}_j,\mathbf{e}_j) = -1$. Dada la métrica $\mathbf{g}$, estos números q y r son los mismos para cualquier punto $p \in M$ y cualquier

base ortonormal de T_pM. La diferencia $q - r$ es la *signatura* de la métrica **g** (y también, por metonimia, la signatura de la variedad $\langle M,\mathbf{g}\rangle$). Obviamente, si la métrica **g** cumple la condición **M3**, $q = n$, $r = 0$, y el concepto de signatura carece de todo interés; por eso nunca se oye hablar de la signatura de una métrica riemanniana propiamente tal. En este libro, siguiendo la práctica de Einstein, definimos las métricas del espaciotiempo de modo que tengan signatura –2 (con $q = 1$ y $r = 3$). En la literatura más reciente es común definirlas en cambio de modo que tengan signatura 2. Desde un punto de vista físico y aun matemático esta diferencia es puramente convencional, pues si **g** es una métrica sobre M con signatura s, la métrica $\mathbf{h} = -\mathbf{g}$ tiene signatura $-s$. Ambas métricas pueden recoger exactamente la misma información, codificándola mediante números que difieren, entre una métrica y la otra, solo por un factor igual a –1.

23 En la nota 23 al capitulo 1 (pp. 206–207) digo cómo hay que definir estos términos si se adopta la convención de signatura preferida hoy. Allí explico también la relación entre sus connotaciones físicas y su empleo en la relatividad especial y general.

24 Como las curvas son aplicaciones lisas, una curva mixta puede siempre descomponerse en segmentos pertenecientes a las distintas clases, posiblemente separados por puntos en que la tangente es de otra clase que en los segmentos adyacentes. Por ejemplo, una curva $\gamma: (0,2) \to M$ puede ser espacialoide sobre el intervalo (0,1), tener una tangente nula en el punto $\gamma(1)$ y ser temporaloide —o nuevamente espacialoide— sobre el intervalo (1,2).

25 Sea γ una curva y $\dot{\gamma}$ un campo vectorial cuyo valor en cada punto del camino de γ es la tangente a γ en ese punto. Si $\nabla\mathbf{g}$ es idéntico a 0, entonces, cualquiera que sea el campo vectorial V, $\nabla\mathbf{g}(V,V,\dot{\gamma}) = \nabla_{\dot{\gamma}}\mathbf{g}(V,V) = 0$, por la ecuación (19). Esto implica que $\mathbf{g}(V,V)$ no varía a lo largo de γ.

26 La idea que subyace a esta definición puede explicarse así. Como sabemos, la métrica **g** determina un producto interno en cada espacio tangente a M. Consideremos un punto $p \in U_x$. Para simplificar la escritura, en esta nota todos los símbolos de campos tensoriales y escalares sobre U_x designarán los valores respectivos de esos campos en p (esto es, **g** significa aquí $\mathbf{g}_p$, g_{ij} significa $g_{ij}(p)$, etc.). Un producto interno **g** sobre T_pM determina un isomorfismo $\varphi: \mathrm{T}_pM \to \mathrm{T}_pM^*$, con $\varphi(V)(W) = \mathbf{g}(V,W)$. Según lo indicado en la p. 162, T_pM se identifica con T_pM^{**} por el isomorfismo natural $\mathbf{i}: \mathrm{T}_pM \to \mathrm{T}_pM^{**}$ definido por $\mathbf{i}(V)(\omega) = \omega(V)$ para todo covector $\omega \in \mathrm{T}_pM^*$. Por lo tanto, la aplicación compuesta $\theta = \mathbf{i} \circ \varphi^{-1}$ es un isomorfismo. Pongamos $\mathbf{g}^*(\alpha,\omega) = \theta(\alpha)(\omega) = \mathbf{i}(\varphi^{-1}(\alpha))(\omega) = \omega(\varphi^{-1}(\alpha)) = \mathbf{g}(\varphi^{-1}(\omega),\varphi^{-1}(\alpha))$, para cualesquiera covectores α, $\omega \in \mathrm{T}_pM^*$. Entonces

—como el lector fácilmente puede probar— $\mathbf{g}^*$ es un producto interno sobre T_pM^*. Como $\mathbf{g} = g_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j$, es claro que la matriz de la aplicación lineal $\varphi{:}\mathrm{T}_pM \to \mathrm{T}_pM^*$ con respecto a las bases $\{\mathrm{d}x^i\}$ y $\{\partial/\partial x^i\}$ es justamente $[g_{ij}]$. Así, $[g_{ij}]^{-1}$ es la matriz de la aplicación inversa φ^{-1} con respecto a las bases mencionadas, y también de θ con respecto a las bases $\{\mathrm{d}x^i\}$ y $\{\mathbf{i}(\partial/\partial x^i)\}$. Por lo tanto, si designamos con g^{ij} al (i,j)-ésimo elemento de $[g_{ij}]^{-1}$, es claro que las ecuaciones (36) en la p. 186 son válidas, y que $\mathbf{g}^* = g^{ij}\partial/\partial x^i \otimes \partial/\partial x^j$ (donde he identificado los $\partial/\partial x^i$ con sus imágenes $\mathbf{i}(\partial/\partial x^i)$). Con la expresión 'la métrica contravariante de $\langle M,\mathbf{g}\rangle$' designo la sección del fibrado de los campos tensoriales de tipo (2,0) sobre M que asigna a cada $p \in M$ el producto interno $\mathbf{g}^*$ sobre T_pM^* que acabo de describir.

27 Sea $\mathscr{V}$ un espacio vectorial n-dimensional real. La función $Q{:}\,\mathscr{V} \to \mathbb{R}$ es una *forma cuadrática* en $\mathscr{V}$ si satisface los requisitos siguientes:

1 $Q(\mathbf{v}) = Q(-\mathbf{v})$, para cualquier $\mathbf{v} \in \mathscr{V}$.
2 Hay una función bilineal $f{:}\,\mathscr{V} \times \mathscr{V} \to \mathbb{R}$, tal que, para todo $\mathbf{v}$, $\mathbf{w} \in \mathscr{V}$,

$$2f(\mathbf{v},\mathbf{w}) = Q(\mathbf{v} + \mathbf{w}) - Q(\mathbf{v}) - Q(\mathbf{w})$$

28 Barrett O'Neill explica así el origen de "la inusual notación ds^2": si M es una variedad riemanniana n-dimensional propiamente tal (la única clase de variedad riemanniana que se conocía cuando la notación se inventó), y p y p' son puntos vecinos de M cuyas respectivas coordenadas son $(x^1,\dots,x^n)$ y $(x^1+\Delta x^1,\dots,x^n+\Delta x^n)$, el vector $\Delta p = \Delta x^i\partial/\partial x^i$ tangente a M en p apunta aproximadamente hacia p'. "Por lo tanto esperamos que el cuadrado de la distancia Δs de p a p' sea aproximadamente igual a

$$|\Delta p|^2 = \mathbf{g}_p(\Delta p,\Delta p) = \Sigma g_{ij}(p)\Delta x^i\Delta x^j$$

tal como en la fórmula $ds^2 = \Sigma g_{ij}dx^idx^j$" (O'Neill 1983, p. 56).

29 La última expresión que doy para R_{ij} depende de la conocida fórmula para la derivada del determinante g de una matriz (g_{ij}) con respecto a uno de sus términos:

$$\frac{\partial g}{\partial g_{ij}} = gg^{ij}$$

donde (g^{ij}) es la matriz inversa, como la definí en la ecuación (36), p. 186. De esto se infiere inmediatamente que

$$\frac{\partial}{\partial g_{ij}} \log g = g^{ij}$$

de donde deducimos, por la ecuación (39), en la p. 187, que

$$\frac{\partial}{\partial x^r} \log g = \left(\frac{\partial}{\partial g_{ij}} \log g \right) \frac{\partial g_{ij}}{\partial x^r} = g^{ij} \frac{\partial g_{ij}}{\partial x^j} = 2\Gamma^i_{ji}$$

Capítulo 8

1 En septiembre de 1909, Max Born presentó a la Asamblea de Naturalistas y Médicos Alemanes reunida en Salzburgo una ponencia en que se refería al movimiento de los cuerpos rígidos en la relatividad especial (Born 1909; cf. Born 1910). En la discusión subsiguiente participó Sommerfeld. Conversando más tarde sobre el mismo asunto, Einstein y Born se percataron de que, al imprimir un movimiento de rotación a un disco rígido, el borde se contraerá en la dirección de cada tangente conforme a la ecuación (1.1) en la p. 22. Esta paradoja lleva ahora el nombre de Ehrenfest, que la comunicó a la *Zeitschrift für Physik* el 29 de septiembre (Ehrenfest 1909). En carta fechada ese mismo día, Einstein le dice a Sommerfeld:

> Me parece muy importante tratar el cuerpo rígido en rotación uniforme con vistas a la extensión del principio de relatividad a sistemas en rotación uniforme según un pensamiento análogo al que intenté desarrollar en [Einstein 1907, §§ 17–18] a propósito de la traslación uniformemente acelerada.
>
> (Einstein, CP, vol. 5, p. 210)

2 En el primero de sus artículos sobre el campo gravitacional estático, Einstein (1912*c*, p. 356) se refiere a un sistema de referencia *K*, con coordenadas *x*, *y*, *z*, en movimiento uniformemente acelerado en la dirección positiva del eje de las *x*. Las distancias se miden en *K* "mediante varas de medir que, comparadas entre sí mientras reposan en un mismo lugar de *K*, poseen la misma longitud"; para longitudes medidas de este modo y también para las relaciones entre las coordenadas *x*, *y*, *z* y otras

longitudes, "han de valer los teoremas de la geometría (*es sollen die Sätze der Geometrie gelten*)". Einstein prosigue:

> No es absolutamente seguro que esta estipulación esté permitida, pues ella contiene postulados físicos que podrían eventualmente probarse erróneos; por ejemplo, es sumamente probable que no valgan en un sistema en rotación uniforme, en el cual, conforme a nuestra definición de las longitudes, la proporción entre la circunferencia y el diámetro tendría que ser diferente de π debido a la contracción de Lorentz.

3 Einstein introduce expresamente esta hipótesis sobre las varas de medir en una nota al pie del texto de sus conferencias de Princeton (1922*c*, p. 39n.; cf. 1956, p. 60n.). Importa señalar que ella no se puede formular exactamente, pues no hay *un único* marco inercial de referencia en que *todas las partes* de un cuerpo acelerado reposen *a la vez* en un momento dado.

4 Sobre la geometría del disco rotatorio se ha escrito muchísimo. Desgraciadamente, gran parte de esa literatura confunde el análisis einsteiniano de la *geometría* establecida mediante varas de medir en el *espacio relativo* de un *marco de referencia* en rotación *uniforme* con la paradoja de Ehrenfest, mencionada en la nota 1, concerniente al comportamiento *mecánico* de un *disco material* que empieza a rotar desde un estado de *reposo*. Nuestro círculo C_2 puede ciertamente representarse mediante un disco material en rotación uniforme —si así nos place— pero no debemos perder de vista que, *por hipótesis*, el borde de *este* disco coincide siempre exactamente con la circunferencia de C_1. Por eso, el disco bajo consideración tiene que suponerse eternamente en movimiento con una velocidad angular constante, o tiene que reemplazarse en la imaginación con un disco material diferente *cada vez* que la velocidad angular cambie. Personalmente. me siento más cómodo con el círculo abstracto evocado en mis explicaciones, que con un disco material como el aquí descrito.

5 Para comprobar estos asertos, calculemos $(x^0(\mathrm{G}) - x^0(\mathrm{P}))$ en función de $(y^0(\mathrm{G}) - y^0(\mathrm{P}))$, y $(x^0(\mathrm{Q}) - x^0(\mathrm{H}))$ en función de $(z^0(\mathrm{Q}) - z^0(\mathrm{H}))$, según el método ilustrado en la ecuación (1), p. 197, recordando que $x^1(\mathrm{P}) = x^1(\mathrm{G}) = x^1(\mathrm{H}) = x^1(\mathrm{Q}) = 0$. Evidentemente, $(x^0(\mathrm{G}) - x^0(\mathrm{P})) = (x^0(\mathrm{Q}) - x^0(\mathrm{H})) = 17{,}4356 \times 0{,}43589 = 7{,}6$.

6 Véase Ciufolini y Wheeler (1995), Will (1993, 1995).

7 Según Balogh y Giampieri (2002), p. 544, "el parámetro η que controla este efecto está actualmente confinado al intervalo $|\eta| \leq 10^{-3}$", esto es, un orden de magnitud mayor que el valor que daban Ciufolini y Wheeler

(1995). Sin embargo, me da la impresión de que Balogh y Giampieri se limitan a repetir el dato de 1976 y que la información de Ciufolini y Wheeler es más reciente.

8 El fragmento se interrumpe aquí.

Obras citadas

Abraham, M. (1912*a*), "Zur Theorie der Gravitation". *Physikalische Zeitschrift,* **13**: 1–4.

Abraham, M. (1912*b*), "Das Elementargesetz der Gravitation". *Physikalische Zeitschrift,* **13**: 4–5.

Abraham, M. (1912*c*), "Berichtigung". *Physikalische Zeitschrift,* **13**: 176.

Abraham, M. (1912*d*), "Der freie Fall". *Physikalische Zeitschrift,* **13**: 310–311.

Abraham, M. (1912*e*), "Die Erhaltung der Energie und der Materie im Schwerkraftfelde". *Physikalische Zeitschrift,* **13**: 311–314.

Abraham, M. (1912*f*), "Relativität und Gravitation: Erwiderung auf eine Bemerkung des Hrn. A. Einsteins". *Annalen der Physik,* **38**: 1056–1058.

Abraham, M. (1912*g*), "Das Gravitationsfeld". *Physikalische Zeitschrift,* **13**: 793–797.

Abraham, M. (1912*h*), "Nochmals Relativität und Gravitation: Bemerkungen zu A. Einsteins Erwiderung". *Annalen der Physik,* **39**: 444–448.

Alpher, R. A. y R. Herman (1948), "Evolution of the universe". *Nature*, **162**: 774.

Alpher, R. A. y R. Herman (1990), "Early work on 'Big-bang' cosmology and the cosmic blackbody radiation". En Bertotti *et al.* (1990), pp. 129–157.

Aristotelis Opera, ex recognitione I. Bekkeris edidit Academia Regia Borussica. Berlin: G. Reimer, 1831. 2 vols. (Mis citas remiten a las páginas, columnas y líneas de esta edición pero proceden de otras ediciones más recientes cuyo texto puede no concordar con el de Bekker.)

Baade, W. (1952), "Extragalactic nebulae. Report to IAU Commission 28". *Transactions of the International Union of Astronomy*, **8**: 397–399.

Balogh A. y G. Giampieri (2002), "Mercury: the planet and its orbit". *Reports on Progress in Physics*, **65**: 529–560.

Barbour, J. B. (1990), "The part played by Mach's Principle in the genesis of relativistic cosmology". En Bertotti *et al.* (1990), pp. 47–66.

Bertotti, B., R. Balbinot, S. Bergia y A. Messina, eds. (1990), *Modern Cosmology in Retrospect.* Cambridge: Cambridge University Press.

Birkhoff, G. D. (1923), *Relativity and Modern Physics.* Cambridge, MA: Harvard University Press.

Blau, S. K. y A. H. Guth, "Inflationary cosmology". En Hawking e Israel (1987), pp. 524–603.

Blumenthal, L. M. (1961), *A Modern View of Geometry*. San Francisco: Freeman.

Boi, L. (1995), *Le problème mathématique de l'espace: Une quête de l'intelligible*. Berlin: Springer.

Bolyai, J. (1832), *Scientia absoluta spatii*. Apéndice de F. Bolyai, *Tentamen juventutem studiosam in elementa matheseos purae elementis ac sublimioris, methodo intuitiva, evidentiaque huic propria, introducendi*, Tomus Primus. Maros Vasarhely: J. et S. Kali.

Boltzmann, L. (1896/98). *Vorlesungen über Gastheorie*. Teil I & II. Leipzig: Barth.

Bondi, H. y T. Gold (1948): "The steady-state theory of the expanding universe". *Monthly Notices of the Royal Astronomical Society*, **108**: 252–270.

Bonola, R. (1945), *Geometrías no euclidianas: exposición histórico-crítica de su desarrollo.* Traducción del italiano por L. Gutiérrez del Arroyo. Buenos Aires: Espasa-Calpe Argentina.

Born, Max (1909), "Über die Dynamik des Elektrons in der Kinematik des Relativitätsprinzips", *Physikalische Zeitschrift*, **10**: 814–817.

Born, Max (1910), "Über die Definition des starren Körpers in der Kinematik des Relativitätsprinzips", *Physikalische Zeitschrift*, **11**: 233–234.

Braginsky, V. B. y V. I. Panov (1972), "Verification of the equivalence of inertial and gravitational mass". *Soviet Physics JETP* **34**: 463–466; Original ruso en *Zh. Eksp. Teor. Fiz.*, **61**: 873–879 (1971).

Carroll, Lewis (1960), *The Annotated Alice: Alice's Adventures in Wonderland & Through the Looking-Glass*, with an Introduction and Notes by Martin Gardner. New York: Bramhall House.

Cayley, A. (1859), “A sixth memoir upon quantics”. *Philosophical Transactions of the Royal Society of London*, **149**: 61–90.

Ciufolini, I. y J. A. Wheeler (1995), *Gravitation and Inertia*. Princeton: Princeton University Press.

Cohen, I. B. (1980), *The Newtonian Revolution: With illustrations of the transformation of scientific ideas.* Cambridge: Cambridge University Press.

Debs, T. A. y M. L. G. Redhead (1996), “The twin ‘paradox’ and the conventionality of simultaneity”. *American Journal of Physics,* **64**: 384–392.

Dicke, R. H., P. J. E. Peebles, P. G. Roll y D. T. Wilkinson (1965), “Cosmic black body radiation”. *Astrophysical Journal*, **142**: 414–419.

Droste, J. (1917), "The field of a single centre in Einstein’s theory of gravitation, and the motion of a particle in that field". *K. Nederlandse Akademie van Wetenschappen, Proceedings*, **A 19**: 197–215. (Comunicado el 27 de mayo de 1916.)

Earman, J. (1995), *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes.* New York: Oxford University Press.

Earman, J. (2001), “Lambda: The constant that refuses to die”. *Archive for History of the Exact Sciences*, **55**: 189–220.

Earman, J. y C. Glymour (1978*a*), “Lost in the tensors: Einstein’s struggles with covariance principles 1912–1916”. *Studies in History and Philosophy of Science,* **9**: 251–278.

Earman, J. y C. Glymour (1978*b*), “Einstein and Hilbert: Two months in the history of general relativity”. *Archive for History of Exact Sciences*, **19**: 291–307.

Eddington, A. S. (1924), *The Mathematical Theory of Relativity,* Second Edition. Cambridge: Cambridge University Press.

Eddington, A. S. (1930), “On the instability of Einstein’s spherical world”. *Monthly Notices of the Royal Astronomical Society*, **90**: 668–678.

Ehresmann, C. (1950),“Les connexions infinitésimales dans un espace fibré différentiable”. *Colloque de Topologie (Espaces Fibrés)*, Bruxelles 1950. Paris: Masson, 1957, pp. 29–55.

Einstein, A. (CP), *The Collected Papers of Albert Einstein.* Princeton: Princeton University Press, 1987– . (Hasta 2002 habían aparecido los vols. 1–7, 8.1 y 8.2.)

Einstein, A. (1905*r*), "Zur Elektrodynamik bewegter Körper". *Annalen der Physik,* **17**: 891–921.

Einstein, A. (1907*j*), "Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen". *Jahrbuch der Radioaktivität und Elektronik,* **4**: 411–462; "Berichtigungen", Ibid., **5**: 98–99 (1908).

Einstein, A. (1911*h*), "Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes". *Annalen der Physik,* **35**: 898–908.

Einstein, A. (1912*c*), "Lichtgeschwindigkeit und Statik des Gravitationsfeldes". *Annalen der Physik,* **38**: 355–369.

Einstein, A. (1912*d*), "Zur Theorie des statischen Gravitationsfeldes". *Annalen der Physik,* **38**: 443–458.

Einstein, A. (1912*h*), "Relativität und Gravitation. Erwiderung auf eine Bemerkung von M. Abraham". *Annalen der Physik,* **38**: 1059–1064.

Einstein, A. (1913*c*), "Zum gegenwärtigen Stande des Gravitationsproblems". *Physikalische Zeitschrift*, **14**: 1249–1262.

Einstein, A. (1915*f*), "Zur allgemeinen Relativitätstheorie". *K. Preußische Akademie der Wissenschaften, Sitzungsberichte*, pp. 778–786.

Einstein, A. (1915*g*), "Zur allgemeinen Relativitätstheorie (Nachtrag)". *K. Preußische Akademie der Wissenschaften, Sitzungsberichte*, pp. 799–801.

Einstein, A. (1915*h*), "Erklärung der Perihelbewegung des Merkurs aus der allgemeinen Relativitätstheorie". *K. Preußische Akademie der Wissenschaften, Sitzungsberichte*, pp. 831–839.

Einstein, A. (1915*i*), "Die Feldgleichungen der Gravitation". *K. Preußische Akademie der Wissenschaften, Sitzungsberichte,* pp. 844–847.

Einstein, A. (1916*e*), "Die Grundlagen der allgemeinen Relativitätstheorie". *Annalen der Physik*, **49**: 769–822.

Einstein, A. (1917*a*), *Über die spezielle und die allgemeine Relativitätstheorie (Gemeinverständich).* Braunschweig: Vieweg.

Einstein, A. (1917*b*), "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie". *K. Preußische Akademie der Wissenschaften, Sitzungsberichte*, pp. 142–152.

Einstein, A. (1918*e*),“Prinzipielles zur allgemeinen Relativitätstheorie”. *Annalen der Physik*, **55**: 241–244.

Einstein, A. (1922*c*), *Vier Vorlesungen über Relativitätstheorie gehalten im Mai 1921 an der Universität Princeton.* Vieweg, Braunschweig.

Einstein, A. (1922), “Bemerkung zu der Arbeit von A. Friedmann, ‘Über die Krümmung des Raumes’”. *Zeitschrift für Physik*, **11**: 326.

Einstein, A. (1923), “Notiz zur der Arbeit von A. Friedmann, ‘Über die Krümmung des Raumes’”. *Zeitschrift für Physik*, **16**: 228.

Einstein, A. (1934), *Mein Weltbild,* Zweite Auflage. Amsterdam: Querido Verlag.

Einstein, A. (1949), “Autobiographisches”. En P. A. Schilpp (1949), pp. 2–94.

Einstein, Albert y J. Laub (1908*a*), “Über die elektromagnetischen Grundgleichungen für bewegte Körper”. *Annalen der Physik,* **26**: 532–540.

Einstein, A. y M. Grossmann (1913), *Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation*. Leipzig: Teubner. Republicado con suplementos en *Zeitschrift für Mathematik und Physik*, **62**: 225–261 (1914).

Einstein, A. y W. de Sitter (1932), “On the relation between the expansion and the mean density of the universe“. *Proceedings of the National Academy of Science*, **18**: 213–214.

Einstein, A., L. Infeld y B. Hoffmann (1938),“The gravitational equations and the problem of motion”. *Annals of Mathematics,* **39**: 65-100.

Einstein, A., H. Born y M. Born (1969), *Briefwechsel 1916–1955.* München: Nymphenburger Verlagshandlung.

Ellis, G. F. R. (1989), “The expanding universe: A history of cosmology from 1917 to 1960”. En D. Howard y J. Stachel (1989), pp. 367–431.

Ellis, G. F. R. (1990), “Innovation, resistance and change: the transition to the expanding universe”. En B. Bertotti *et al*. (1990), pp. 98–113.

Eötvös, R. von (1889), “Über die Anziehung der Erde auf verschiedene Substanzen”. En Eötvös (1953), *Gesammelte Arbeiten*, Budapest: Akadémiai Kiado, pp. 17–20. (Esta comunicación fue leída a la Academia Húngara de Ciencias el 20 de enero de 1889 y publicada luego en *Mathematische und naturwissenschaftliche Berichte Ungarns*, vol. 8, pp. 65–68.)

Eötvös, R., D. Pekar y E. Fekete (1922), "Beiträge zum Gesetze der Proportionalität von Trägheit und Gravität". *Annalen der Physik*, **68**: 11–66. (Los resultados de esta investigación fueron comunicados a la Academia Húngara de Ciencias en 1909.)

Euclides, *Elementa,* edidit I. L. Heiberg. Leipzig: B. G. Teubner. 5 vols.

Feynman, E. F. B. Morinigo y W. G. Wagner (1995), *Feynman Lectures on Gravitation*, edited by B. Hatfield, with a Foreword by J. Preskill and K. S. Thorne. Reading, MA: Addison-Wesley. (Feynman dio este curso en 1962/63).

Frege, G. (KS), *Kleine Schriften*, herausgegeben von I. Angelelli. Darmstadt: Wissenschaftliche Buchgesellschaft, 1967.

Freudenthal, H. (1957), "Zur Geschichte der Grundlagen der Geometrie". *Nieuw Archief vor Wiskunde*, **5**: 105–142.

Freudenthal, H. (1960), "Die Grundlagen der Geometrie um die Wende des 19. Jahrhunderts". *Mathematisch-physikalische Semesterbericht*, **7**: 2–25.

Friedmann, A. (1922), "Über die Krümmung des Raumes". *Zeitschrift für Physik*, **10**: 377–386.

Friedmann, A. (1924), "Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes". *Zeitschrift für Physik*, **21**: 326–332.

Galileo Galilei (EN), *Le Opere,* Nuova ristampa della Edizione Nazionale. Firenze: G. Barbera, 1964–1966. 20 vols.

Galileo Galilei (1623), *Il saggiatore nel quale con bilancia esquisita e giusta si ponderano le cose contenute nella Libra astronomica e filosofica di Lotario Sarsi Sigensano, scritto in forma di lettera all' ill.^{mo} et rever.^{mo} mons.^{re} d. Virginio Cesarini*. Roma: Giacomo Mascardi.

Galileo Galilei (1632), *Dialogo . . . sopra i due Massimi Sistemi del Mondo Tolemaico, e Copernicano; proponendo indeterminatamente le ragioni Filosofiche, e Naturali tanto per l'una, quanto per l'altra parte*. Firenze: Gio.Batista Landini.

Garnavich, P. M., R. P. Kirshner, P. Challis, J. Tonry, R. L. Gilliland, R. C. Smith, A. Clocchiatti, A. Diercks, A. V. Filippenko, M. Hamuy, C. J. Hogan, B. Leibundgut, M. M. Phillips, D. Reiss, A. G. Riess, B. P. Schmidt, J. Spyromilio, C. Stubbs, N. B. Suntzeff y L. Wells (1998), "Constraints on Cosmological Models from Hubble Space

Telescope Observations of High-z Supernovae". *Astrophysical Journal*, **493**: L53–57.

Gauss, C. F. (WW), *Werke*, Achter Band, herausgegeben von der K. Gesellschaft der Wissenschaften zu Göttingen. Leipzig: B. G. Teubner, 1900.

Gauss, C. F. (1828), “Disquisitiones generales circa superficies curvas”. *Commentationes societatis regiae scientiarum Gottingensis Recentiores*, **6**: 99–146.

Geroch, Robert P. y P. S. Jang (1975), “Motion of a body in General Relativity”. *Journal of Mathematical Physics*, **16**: 65–67.

Göckeler, M. y T. Schücker (1987), *Differential Geometry, Gauge Theories, and Gravity.* Cambridge: Cambridge University Press.

Gödel, K. (1949), “An example of a new type of cosmological solutions of Einstein’s field equations of gravitation”. *Reviews of Modern Physics,* **21**: 447–450.

Gödel, K. (1949*a*), “A remark about the relationship between relativity theory and idealistic philosophy”. En P. A. Schilpp (1949), pp. 555–562.

Gödel, K. (1949*b*), “Lecture on rotating universes”. En Gödel (1995), *Collected Works*, Volume III, Unpublished essays and lectures. New York: Oxford University Press, pp. 269–287.

Gödel, K. (1952), “Rotating universes in general relativity theory”. En *Proceedings of the International Congress of Mathematicians; Cambridge, Massachussetts, U.S.A. August 30–September 6, 1950.* Providence RI: American Mathematical Society, vol. I, pp. 175–181.

Guth, A. H. (1982), “Inflationary universe: A possible solution of the horizon and flatness problems”. *Physical Review* **D23**: 347–356.

Hafele, J. C. y R. E. Keating (1972), “Around-the-world atomic clocks”. *Science,* **177**: 166–170.

Halmos, P. R. (1958), *Finite-dimensional Vector Spaces.* Second Edition. Princeton, NJ: Van Nostrand.

Hall, A. R. y M. B. Hall, eds. (1962), *Unpublished Science Papers of Isaac Newton.* A Selection from the Portsmouth Collection in the University Library, Cambridge. Cambridge: Cambridge University Press.

Hawking, S. W. (1969), “The existence of cosmic time functions”. *Royal Society of London Proceedings,* **A308**: 433–435.

Hawking, S. W. y G. F. R. Ellis (1973), *The Large Scale Structure of Space-Time.* Cambridge: Cambridge University Press.

Hawking, S. W. y W. Israel, eds. (1987), *Three hundred years of gravitation*. Cambridge: Cambridge University Press.

Healey, R. (2002), Review of Jeremy Butterfield (ed.), *The Arguments of Time*, Oxford: Clarendon Press, 1999. *British Journal for the Philosophy of Science*, **53**: 459–463.

Heath, T. L. (1956), *The Thirteen Books of Euclid's Elements.* Translated from the text of Heiberg with introduction and commentary. Second edition, revised with additions. New York: Dover. 3 vols.

Hilbert, D. (1899), "Die Grundlagen der Geometrie," en *Festschrift zur Feier der Enthüllung des Gauss-Weber Denkmals*, Leipzig: B. G. Teubner, pp. 3–92.

Howard, D. y J. Stachel (1989), *Einstein and the History of General Relativity*, Boston: Birkhäuser (*Einstein Studies*, vol. 1).

Hoyle, F. (1948), "A new model for the expanding universe". *Monthly Notices of the Royal Astronomical Society*, **108**: 372–382.

Hubble, E. P. (1929), "A relation between distance and radial velocity among extragalactic nebulae". *Proceedings of the National Academy of Sciences,* **15**: 168–173.

Jeffreys, H. (1919), "On the Crucial Test of Einstein's Theory of Gravitation". *Monthly Notices of the Royal Astronomical Society*, **80**: 138–154.

Kaiser, D. (1998), "A ψ is just a ψ? Pedagogy, practice and the reconstitution of General Relativity, 1942–1975". *Studies in History and Philosophy of Modern Physics*, **29B**: 321–338.

Klein, F. (1871), "Über die sogenannte Nicht-Euklidische Geometrie". *Mathematische Annalen*, **4**: 573–625.

Klein, F. (1872), *Vergleichende Betrachtungen über neuere geometrische Forschungen.* Erlangen: A. Duchert.

Klein, F. (1873), "Über die sogenannte Nicht-Euklidische Geometrie (Zweiter Aufsatz)," *Mathematische Annalen,* **6**: 112–145.

Klein, F. (1893), "Vergleichende Betrachtungen über neuere geometrische Forschungen", *Mathematische Annalen,* **43**: 63–100. (Version corregida de Klein 1872.)

Klein, F. (1911),"Über die geometrischen Grundlagen der Lorentz-Gruppe", *Physikalische Zeitschrift*, **12**: 17–27.

Kobayashi, S. y K. Nomizu (1963/1969), *Foundations of Differential Geometry.* New York: Wiley. 2 vols.

Kretschmann, E. (1917), "Über den physikalischen Sinn der Relativitäts-postulaten". *Annalen der Physik*, **53**: 575–614.

Lagrange, J. L. (1788). *Mécanique analitique.* Paris: Chez la Veuve Desaint.

Lange, L. (1885), "Über das Beharrungsgesetz". *K. Sächsische Gesellschaft der Wissenschaften zu Leipzig, Berichte, Math.-Phys. Classe*, **37**: 333–351.

Laue, M. von (1911*a*), "Zur Dynamik der Relativitätstheorie". *Annalen der Physik,* **35**: 524–542.

Laue, M. von (1911*b*), "Zur Diskussion über den starren Körper in der Relativitätstheorie". *Physikalische Zeitschrift*, **12**: 85–87.

Laue, M. von (1911*c*), *Das Relativitätsprinzip.* Braunschweig: Vieweg.

Lemaître, G. (1927), "Un univers homogène de masse constante et de rayon croissant, rendant compte de la vitesse radiale de nébuleuses extra-galactiques". *Annales de la Sociéte Scientifique de Bruxelles*, **47A**: 49–59.

Levi-Civita, T. (1917), "Nozione di parallelismo in una varietà qualunque". *Rendiconti del Circolo Matematico di Palermo*, **42**: 173–205.

Linde, A. (1983), "Chaotic inflation". *Physical Letters*, **B129**: 177–181.

Linde, A. (1986), "Eternal chaotic inflation". *Modern Physics Letters,* **A1**: 81.

Linde, A. (1987), "Inflation and quantum cosmology". En Hawking e Israel (1987), pp. 604–630.

Lindquist R. W. y J. A. Wheeler (1957), "Dynamics of a lattice universe by the Schwarzschild cell method", *Reviews of Modern Physics*, **29**: 432–433.

Lobachevsky, N. I. (1837), "Géométrie imaginaire". *Journal für die reine und angewandte Mathematik,* **17**: 295–320.

Lobachevsky, N. I. (1840), *Geometrische Untersuchungen zur Theorie der Parallellinien*. Berlin: F. Fincke.

Lobachevsky, N. I. (1856), *Pangéométrie ou précis de géométrie fondée sur une théorie générale et rigoureuse des parallèles*. Kazan: Universitet.

[Locke, J.] (1690), *An Essay concerning Humane Understanding*, in four books. London: Printed for Thomas Basset and sold by Edward Mory. (Anónimo; la segunda edición se publicó con el nombre del autor.)

Lorentz, H. A. (CP), Co*llected Papers*, edited by P. Zeeman and A. D. Fokker. The Hague: Nijhoff, 1935–1939. 9 vols.

Lorentz, H. A. (1895), *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern.* Leiden: Brill. (Reproducido en Lorentz (CP), vol. 5, pp. 1–137).

Lorentz, H. A. (1899), "Theorie simplifiée des phénomènes électriques et optiques dans les corps en mouvement". En Lorentz CP, vol. 5, pp. 139–155. (Publicado originalmente en holandés en *Versl. K. Nederlandse Akademie van Wetenschappen,* **7**: 507–522).

Mach, E. (1883), *Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt*. Leipzig: F. A. Brockhaus.

Malament, D. (1977), "Causal theories of time and the conventionality of simultaneity". *Noûs*, **11**: 293–300.

Mayo, D. G. (1996), *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.

McVittie, G. C. (1933), "The mass-particle in an expanding universe". *Monthly Notices of the Royal Astronomical Society*, **93**: 325–339.

Michelson, A. A. (1881), "The relative motion of the earth and the luminiferous ether". *American Journal of Science,* **22**: 120–129.

Michelson, A. A. y E. W. Morley (1886), "Influence of motion of the medium on the velocity of light". *American Journal of Science*, **31**: 377–386.

Michelson, A. A. y E. W. Morley (1887), "On the relative motion of the earth and the luminiferous aether". *Philosophical Magazine,* **24**: 449–463.

Minkowski, H. (1908), "Die Grundgleichungen fur die elektromagnetischen Vorgänge in bewegten Körper". *Göttinger Nachrichten,* pp. 53–111.

Minkowski, H. (1909), "Raum und Zeit". *Physikalische Zeitschrift*, **10**: 104–111.

Minkowski, H. (1915), "Das Relativitätsprinzip". *Jahresbericht der Deutschen Mathematiker-Vereinigung,* **24**: 372–382. (Conferencia

dictada el 5 de noviembre de 1907 ante la Sociedad Matemática de Göttingen.)

Misner, C. W., K. S. Thorne y J. A. Wheeler (1973), *Gravitation.* San Francisco: W. H. Freeman.

Mosterín, J. y R. Torretti (2002), *Diccionario de lógica y filosofía de la ciencia.* Madrid: Alianza Editorial.

Nagel, E. (1939), "The formation of modern conceptions of formal logic in the development of geometry". *Osiris*, **7**: 142–224.

Neumann, C. (1870), *Über die Principien der Galilei-Newton'schen Theorie*, Akademische Antrittsvorlesung gehalten in der Aula der Universität Leipzig am 3. November 1869. Leipzig: Teubner.

Newton, I. (*Principia*), *Philosophiæ Naturalis Principia Mathematica.* The Third Edition (1726) with variant readings assembled and edited by A. Koyré and I. B. Cohen with the Assistance of A. Whitman. Cambridge, MA: Harvard University Press. 2 vols.

Newton, I. (1726), *Philosophiæ naturalis principia mathematica,* editio tertia aucta & emendata. Londini: Apud Guil. & Joh. Innys, Regiæ Societatis typographos.

Nomizu, K. (1954), "Invariant affine connections on homogeneous spaces". *American Journal of Mathematics*, **76**: 33–65.

Norton, J. (1984), "How Einstein found his Field Equations: 1912–1915?". *Historical Studies in the Physical Sciences,* **14**: 253–316.

Norton, J. (1985), "What was Einstein's Equivalence Principle?". *Studies in History and Philosophy of Science,* **16**: 203–246.

O'Neill, B. (1983), *Semi-Riemannian Geometry, with Applications to Relativity*. New York: Academic Press.

Pasch, M. (1882), *Vorlesungen über neuere Geometrie.* Leipzig: Teubner.

Pauri, M. (1991), "The universe as a scientific object". En E. Agazzi y A. Cordero, eds., *Philosophy and the Origin and Evolution of the Universe*, Dordrecht: Kluwer, pp. 291–339.

Peebles, P. J. E. (1993), *Principles of Physical Cosmology.* Princeton, NJ: Princeton University Press.

Penrose, R. (1968), "Structure of Space-Time". En C. DeWitt y J. A. Wheeler, eds., *Battelle Rencontres*: *1967 Lectures in Mathematics and Physics*. New York: Benjamin, pp. 121–235.

Penzias, A. A. y R. W. Wilson (1965), "A measurement of excess antenna temperature at 4080 Mc/s", *Astrophysical Journal*, **142**: 419–421.

Perlmutter, S., G. Aldering, M. Della Valle, S. Deustua, R. S. Ellis, S. Fabbro, A. Fruchter, G. Goldhaber, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, R. A. Knop, C. Lidman, R. G. McMahon, Peter Nugent, R. Pain, N. Panagia, C. R. Pennypacker, P. Ruiz-Lapuente, B. Schaefer y N. Walton (1998), "Discovery of a Supernova Explosion at Half the Age of the Universe and its Cosmological Implications". *Nature*, **391**: 51–54.

Planck, M. (1907), "Zur Dynamik bewegter Systeme". *K. Preußische Akademie der Wissenschaften*, *Sitzungsberichte,* pp. 542–570.

Poincaré, H. (VS), *La valeur de la science*. Paris: Flammarion, 1970. (Primera edición: Paris, Flammarion, 1904).

Poincaré, H. (1887), "Sur les hipothèses fondamentales de la géométrie," *Bulletin de la Société mathématique de France*, **15**: 203–216.

Poincaré, H. (1898), "La mésure du temps". *Revue de métaphysique et de morale,* **6**: 1–13.

Poincaré, H. (1902), *La science et l'hypothèse.* Paris: Flammarion.

Poincaré, H. (1906), "Sur la dynamique de l'électron". *Rendiconti del Circolo matematico di Palermo*, **21**: 129–175. Resumen en *Comptes rendus de l'Académie des Sciences,* **140** (1905): 1504–1508.

Poncelet, J. V. (1822), *Traité des propriétés projectives des figures.* Paris: Bachelier.

Redhead, M. (1993), "The conventionality of simultaneity". En J. Earman, A. I. Janis, G. J. Massey y N. Rescher, eds., *Philosophical Problems of Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*, Pittsburgh: University of Pittsburgh Press, pp. 103–128.

Reichenbach, H. (1924), *Axiomatik der relativistischen Raum-Zeit-Lehre.* Braunschweig: Vieweg.

Reichenbach, H. (1928), *Philosophie der Raum-Zeit-Lehre.* Berlin: W. de Gruyter.

Ricci, G. y T. Levi-Civita (1901), "Méthodes de calcul différentiel absolu et leurs applications". *Mathematische Annalen,* **54**: 125–201.

Riemann, B. (1854), "Über die Hypothesen, welche der Geometrie zugrunde liegen". *Abhandlungen der Kgl. Gesellschaft der*

Wissenschaften zu Göttingen, **13**: 133–152 (1867). (Lección inaugural dictada en 1854, publicada póstumamente por R. Dedekind en 1967; traducción castellana por R. Torretti en *Diálogos*, **31**: 151–168 (1978), y por J. Ferreirós en *Riemanniana selecta*, Madrid: Consejo Superior de Investigaciones Científicas (2000); traducción inglesa por M. Spivak en Spivak (1979), vol. 2, pp. 135 y ss.)

Riemann, B. (1861), "Commentatio mathematica, qua respondere tentatur quaestioni ab illustrissima Acad. Parisiensi propositae". En *Bernhard Riemanns gesammelte mathematische Werke und wissenschaftlicher Nachlass*, Leipzig: Teubner, 1876, pp. 391–404.

Rindler, W. (1956), "Visual horizons in world models". *Monthly Notices of the Royal Astronomical Society,* **116**: 662–677.

Robertson, H. P. (1929), "On the foundations of relativistic cosmology". *National Academy of Sciences Proceedings,* **15**: 822–829.

Robertson, H. P. (1933), "Relativistic cosmology", *Reviews of Modern Physics*, **5**: 62–90.

Robertson, H. P. (1935/36), "Kinematics and world-structure", *Astrophysical Journal*, **82**: 284–301; **83**: 187–201, 257–271.

Robertson, H. P. y T. W. Noonan (1969), *Relativity and Cosmology.* Philadelphia: W. B. Saunders Company.

Roll, P. G., R. Krotkov y R. H. Dicke (1964): "The equivalence of inertial and passive gravitational mass". *Annals of Physics*, **26**: 442–517.

Rosenfeld, B. A. (1988), *A History of Non-Euclidean Geometry: Evolution of the Concept of a Geometric Space*, translated [from the Russian] by Abe Shenitzer. New York: Springer.

Rossi, B. y D. B. Hall (1941), "Variation of the rate of decay of mesotrons with momentum". *Physical Review,* **59**: 223–228.

Russell, B. (1897), *An Essay on the Foundations of Geometry.* Cambridge: Cambridge University Press.

Saccheri, G. (1733), *Euclides ab omni nævo vindicatus sive conatus geometricus quo stabiliuntur prima ipsa universæ geometriæ principia.* Mediolani: Ex Typographia Pauli Antonii Montani.

Sarkar, S. y J. Stachel (1999), "Did Malament Prove the Non-Conventionality of Simultaneity in the Special Theory of Relativity?" *Philosophy of Science*, **6**: 208–220.

Schilpp, P. A., ed. (1949), *Albert Einstein: Philosopher-Scientist*. Evanston IL: Library of Living Philosophers.

Schouten, J. A. y E. R. van Kampen (1930), "Zur Einbettungs- und Krümmungstheorie nichtholonomer Gebilde". *Mathematische Annalen,* **103**: 752–783.

Schur, F. (1886), "Über den Zusammenhang der Räume constanten Riemann'schen Krümmungsmaasses mit den projectiven Räumen". *Mathematische Annalen*, **27**: 537–567.

Schutz, B. F. (1980), *Geometrical Methods of Mathematical Physics*. Cambridge: Cambridge University Press.

Schwarzschild, K. (1916*a*), "Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie". *K. Preußische Akademie der Wissenschaften, Sitzungsberichte*, pp. 189–196. (Comunicado el 13 de enero de 1916.)

Schwarzschild, K. (1916*b*), "Über das Gravitationsfeld einer Kugel aus inkompressibler Flüssigkeit nach der Einsteinschen Theorie". *K. Preußische Akademie der Wissenschaften, Sitzungsberichte*, pp. 424–434.

Schweber, S. S. (1994), *QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga*. Princeton: Princeton University Press, 1994.

Schwinger, J., ed. (1958), *Selected Papers on Quantum Electrodynamics*. New York: Dover.

Sitter, W. de (1917*a*), "On the relativity of inertia. Remarks concerning Einstein's latest hypothesis". *K. Nederlandse Akademie van Wetenschappen, Proceedings*, **19**: 1217–1225.

Sitter, W. de (1917*b*), "On the curvature of space". *K. Nederlandse Akademie van Wetenschappen, Proceedings*, **20**: 229–243.

Sitter, W. de (1917*c*), "On Einstein's theory of gravitation, and its astronomical consequences. III". *Monthly Notices of the Royal Astronomical Society*, **78**: 3–28.

Smith, R. (1982), *The Expanding Universe: Astronomy's 'Great Debate' 1900–1931*. Cambridge: Cambridge University Press.

Sommerfeld, A. (1910*a*), "Zur Relativitätstheorie. I. Vierdimensionale Vektoralgebra". *Annalen der Physik,* **32**: 749–776.

Sommerfeld, A. (1910*b*), "Zur Relativitätstheorie. II. Vierdimensionales Vektoranalysis". *Annalen der Physik,* **33**: 649–689.

Spivak, M. (1965), *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus.* New York: W. A. Benjamin. (Hay versión española por G. P. Xufré: M. Spivak, *Cálculo en variedades*, Barcelona, Reverté, 1972).

Spivak, M. (1979), *A Comprehensive Introduction to Differential Geometry*, second edition. Berkeley: Publish or Perish. 5 vols.

Stachel, J. (1980), "Einstein and the rigidly rotating disk". En A. Held, ed. (1980), *General Relativity and Gravitation: One Hundred Years after the Birth of Albert Einstein*, New York: Plenum Press, vol. 1, pp. 1–15.

Stachel, J. (1989), "Einstein's search for general covariance, 1912–1915". En D. Howard y J. Stachel (1989), pp. 63–100.

Stein, H. (1991), "On Relativity Theory and the openness of the future". *Philosophy of Science,* **58**: 147–167.

Thomson, J. (1884), "On the Law of Inertia, the Principle of Chronometry and the Principle of Absolute Clinural Rest, and of Absolute Rotation". *Royal Society of Edinburgh Proceedings,* **12**: 568–578.

Tolman, R. C. (1934), *Relativity, Thermodynamics and Cosmology.* Oxford: Clarendon Press.

Torretti, R. (1978), *Philosophy of Geometry from Riemann to Poincaré.* Dordrecht: Reidel.

Torretti, R. (1994), *La geometría del universo y otros ensayos de filosofía natural.* Mérida, Venezuela: Universidad de los Andes.

Torretti, R. (1996), *Relativity and Geometry*. New York: Dover. (Reimpresión corregida de Torretti, *Relativity and Geometry*, Oxford: Pergamon, 1983).

Torretti, R. (1999), *The Philosophy of Physics*. New York: Cambridge University Press.

Veblen, O. y J. H. C. Whitehead (1932), *The Foundations of Differential Geometry.* Cambridge University Press, Cambridge.

Vizgin, V. (2002), "The role played by Mach's ideas in the genesis of the General Theory of Relativity". En Y. Balashov y V. Vizgin, eds., *Einstein Studies in Russia*, Boston: Birkhäuser (*Einstein Studies*, vol. 10), pp. 45–89.

Wald, R. M. (1984), *General Relativity*. Chicago: University of Chicago Press

Walker, A. G. (1935*a*), "On Riemannian spaces with spherical symmetry about a line and the conditions for isotropy in general relativity". *Quarterly Journal of Mathematics*, **6**: 81–93.

Walker. A. G. (1935*b*), "On the formal comparison of Milne's kinematical system with the systems of general relativity". *Monthly Notices of the Royal Astronomical Society,* **95**: 263–269.

Walker, A. G. (1937), "On Milne's theory of world-structure". *Proceedings of the London Mathematical Society*, **42**: 90–127.

Weinberg, S. (1972), *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: Wiley.

Westfall, R. S. (1980), *Never at rest: A Biography of Isaac Newton.* Cambridge: Cambridge University Press.

Whittaker, E. T. (1928), "Note on the law that light-rays are the null-geodesics of a gravitational field". *Cambridge Philosophical Society Proceedings*, pp. 32–34.

Whittaker, E. T. (1951/53), *A History of the Theories of Aether and Electricity.* London: Thomas Nelson & Sons.

Will, C. W. (1993), *Theory and Experiment in Gravitational Physics*, revised edition. Cambridge: Cambridge University Press.

Will, C. W. (1995). "The confrontation between General Relativity and experiment: A 1995 update". En Hall, G. S. y J. R. Pulham, eds., *General Relativity*, Proceedings of the Forty Sixth Scottish Universities Summer School in Physics; Aberdeen, July 1995. A NATO Advanced Study Institute, pp. 239–281.

Winnie, J. A. (1970), "Special relativity without one-way velocity assumptions". *Philosophy of Science,* **37**: 81–99, 223–238.

Winnie, J. W. (1986), "Invariants and objectivity: A theory with applications to relativity and geometry". En R. G. Colodny, ed., *From Quarks to Quasars*, Pittsburgh: Pittsburgh University Press, pp. 71–180.

Recomiendo calurosamente este magnífico texto de estudio, que apareció mientras el presente libro estaba en prensa:

Hartle, James B. (2003), *Gravity: An Introduction to Einstein's General Relativity.* San Francisco: Addison-Wesley.

Origen de los capítulos

[1] Relatividad y espaciotiempo
Inédito. Una versión algo más breve que esta y sin notas debería aparecer eventualmente en el volumen a cargo de Anna Estany que la *Enciclopedia Iberoamericana de Filosofía* dedica a la filosofía de las ciencias exactas, naturales y sociales.

[2] El orden del tiempo
Traducción de Torretti, "On Relativity, Time-Reckoning, and the Topology of Time Series", en Jeremy Butterfield, ed., *The Arguments of Time*. Oxford: Published for the British Academy by Oxford University Press, 1999, pp. 66–82.

[3] Gravedad y curvatura
Traducción de Torretti, "Gravity as Spacetime Curvature". *Physics in Perspective*, **2**: 118–134 (2000).

[4] Modelos del mundo
Traducción de Torretti, "Spacetime Models for the World". *Studies in the History and Philosophy of Modern Physics*, **31**: 171–186 (2000) (Copyright © 2000, Elsevier Science). La traducción fue autorizada, mas no examinada, por Elsevier Science.

[5] Los universos de Friedmann: derivando la geometría
Traducción de Torretti, "The Friedmann Worlds: Deriving the geometry". Manuscrito repartido a los asistentes al curso *¿Se acelera la expansión del Universo? Cosmología y Filosofía*, ofrecido por la Universidad Internacional Meléndez Pelayo (Santander, España) en julio de 2000.

[6] Geometría decimonónica
Traducción de Torretti, "Geometry, Nineteenth Century". Artículo en *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/geometry-19th/.

[7] **Apéndice matemático**
Traducción del apéndice terminológico de [5].

Índice

Términos matemáticos explicados en el capítulo 7

Términos comentados en los capítulos 1–6 y 8 y nombres propios de personas

Se terminó de imprimir
en los talleres digitales de RIL® editores
en el mes de mayo de 2003. Santiago de Chile